





PT New nucleic acid encoding plant sugar-transport proteins, useful for  
 PT preparing transgenic plants with altered carbohydrate distribution.  
 XX  
 PS  
 XX Example 4; Fig 2; 54pp; English.  
 CC The present invention relates to the isolation of plant polynucleotide  
 CC sequences encoding an Arabidopsis thaliana-like sugar transport protein  
 CC or Beta vulgaris-like sugar transport protein. The polynucleotide  
 CC sequences are useful for altering the level of sugar transport proteins  
 CC in plants, i.e. for control of carbohydrate transport and distribution in  
 CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,  
 CC rice, soybeans, and wheat), and, for studying carbohydrate flows and  
 CC sugar transport. The polynucleotide sequences can also be used to isolate  
 CC cDNA sequences and genes that encode homologues of the new proteins. The  
 CC present sequence represents a soybean Beta vulgaris-like sugar transport  
 CC protein  
 XX  
 SQ Sequence 523 AA;

Query Match 60.3%; Score 1542; DB 5; Length 523;  
 Best Local Similarity 62.0%; Pred. No. 1.8e-146;  
 Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKNVYASICAILASMASVILGYDVGMSGAAMYIK 50  
 Db 1 MTEGKLVAAEAHKTLDQDFPKKRNKYAFACAMLASMTSILLGYDVGMSGAAMYIK 60

QY 51 KDLNITDVQLIELIGLSLFSFGAGARTSDRIGRLTVFAAVIFFVGSLLMGFAVN 110  
 Db 61 RDLKVSDEQIEILLGIINLYSLIGSLAGRTSDWIGPRYTVFAGTFFVFGALLMGFSPN 120

QY 111 YGMLMAGRFVAGVGVGGMIAPVYTAIEISPAASRGFLTTPPEVFNIGILLGSLNPAF 170  
 Db 121 YSFLMFGRFVAGIGIGVLMIAPVYTAIEVSPASRGFLTTPPEVFNIGILLGSLNPAF 180

QY 171 ARPLHLGWRVYMLAIGAVPSGLLALLVFCMPESPRMLVKGLRADARAVLEKTSATPEEA 230  
 Db 181 SKLTLKVGWRMMLGVGAIPSVLLTVGVLMAMPESPRMLVMRGLGEARKVNLKTSDEEA 240

QY 231 AERLADIKAAAGIPKGLDGDVTVPGKEGGELQVWKLLISPTPAVRRLLSAVGLHF 290  
 Db 241 QLRLAEIKQAAGIPESCNDDVQVKNOSNGEG---VWKELFLYPTPAIRHIVIAALGIHF 297

QY 291 FQOASGSDSVVQYSARLFKSGAGITDDNKLGVTCVAGVTKTFFILVATFLLDRAGRPL 350  
 Db 298 FQOASGVDVAVLYSPRIFERAGITNDTHKLLATVAVGVKTVFILATFLLDRAGRPL 357

QY 351 LISTGGMIVSLICLGSGLTVAGHHPTDKVAVAVLCIASTLSYIAFFSIGLPITGVYTS 410  
 Db 358 LSSVGGMVLSTLTLAISLTVI-DHSERKLMWAVGSSIAMVLAIVATFSGAGPITWYSS 416

QY 411 EIPPLQVRALGFVGVASNRVTSNAVISMTELSLSKAITTGGSFYSGIAAVAWVFFTC 470  
 Db 417 EIPPLRUAQAAAGVAVNTTSAVVSMTFLSUTRAITIGGAFPLYCGIATVGVWIFFYTV 476

QY 471 LPETGRGLEMGKLF 487  
 Db 477 LPETGRKLEDMGSG 493

RESULT 10  
 ABU08337  
 ID ABU08337 standard; protein; 523 AA.  
 XX  
 AC ABU08337;  
 XX  
 DT 29-MAY-2003 (first entry)  
 XX  
 DE Soybean sugar transport protein #3.  
 XX  
 DE Beta vulgaris-like sugar transport protein; corn; rice; wheat;  
 KW plant sugar transport protein; carbohydrate transport; soybean;  
 KW carbohydrate distribution; plant.

XX Glycine max.  
 OS US2002178468-A1.  
 XX  
 PD 28-NOV-2002.  
 XX  
 XX 17-JAN-2002; 2002US-00051902.  
 PF  
 XX 24-APR-1998; 98US-0083044P.  
 PR  
 PR 14-APR-1999; 99US-00291922.  
 XX  
 PA (ALLE/) ALLEN S M.  
 PA (HITZ/) HITZ W D.  
 PA (KINN/) KINNEY A J.  
 PA (TING/) TINGEY S V.  
 XX  
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;  
 XX  
 XX WPI; 2003-340957/32.  
 DR N-PSDB; ABX93209.  
 DR  
 XX  
 PT Novel plant sugar transport proteins and nucleic acid encoding the  
 PT protein useful for producing transgenic plants having altered levels of  
 PT sugar transport protein.  
 XX  
 PS Claim 10; Fig 2; 56pp; English.  
 XX  
 CC The present invention relates to the isolation of Arabidopsis thaliana-  
 CC like or Beta vulgaris-like sugar transport proteins, and the  
 CC polynucleotide sequences encoding them. The plant sugar transport  
 CC proteins of the invention have been isolated from corn, rice, soybean,  
 CC and wheat. The polypeptides of the invention may be used for altering the  
 CC level of expression of a sugar transport protein in a host cell, by  
 CC transforming a host cell with a chimeric construct encoding all, or a  
 CC portion of the sugar transport protein, in sense or antisense  
 CC orientation. Particularly, the polypeptides may provide a means to  
 CC control carbohydrate transport and distribution in plants. ABU08334-  
 CC ABU08339 represent Beta vulgaris-like sugar transport proteins  
 XX  
 SQ Sequence 523 AA;

Query Match 60.3%; Score 1542; DB 6; Length 523;  
 Best Local Similarity 62.0%; Pred. No. 1.8e-146;  
 Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKNVYASICAILASMASVILGYDVGMSGAAMYIK 50  
 Db 1 MTEGKLVAAEAHKTLDQDFPKKRNKYAFACAMLASMTSILLGYDVGMSGAAMYIK 60

QY 51 KDLNITDVQLIELIGLSLFSFGAGARTSDRIGRLTVFAAVIFFVGSLLMGFAVN 110  
 Db 61 RDLKVSDEQIEILLGIINLYSLIGSLAGRTSDWIGPRYTVFAGTFFVFGALLMGFSPN 120

QY 111 YGMLMAGRFVAGVGVGGMIAPVYTAIEISPAASRGFLTTPPEVFNIGILLGSLNPAF 170  
 Db 121 YSFLMFGRFVAGIGIGVLMIAPVYTAIEVSPASRGFLTTPPEVFNIGILLGSLNPAF 180

QY 171 ARPLHLGWRVYMLAIGAVPSGLLALLVFCMPESPRMLVKGLRADARAVLEKTSATPEEA 230  
 Db 181 SKLTLKVGWRMMLGVGAIPSVLLTVGVLMAMPESPRMLVMRGLGEARKVNLKTSDEEA 240

QY 231 AERLADIKAAAGIPKGLDGDVTVPGKEGGELQVWKLLISPTPAVRRLLSAVGLHF 290  
 Db 241 QLRLAEIKQAAGIPESCNDDVQVKNOSNGEG---VWKELFLYPTPAIRHIVIAALGIHF 297

QY 291 FQOASGSDSVVQYSARLFKSGAGITDDNKLGVTCVAGVTKTFFILVATFLLDRAGRPL 350  
 Db 298 FQOASGVDVAVLYSPRIFERAGITNDTHKLLATVAVGVKTVFILATFLLDRAGRPL 357

QY 351 LISTGGMIVSLICLGSGLTVAGHHPTDKVAVAVLCIASTLSYIAFFSIGLPITGVYTS 410  
 Db 358 LSSVGGMVLSTLTLAISLTVI-DHSERKLMWAVGSSIAMVLAIVATFSGAGPITWYSS 416



PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138550P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139444P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140033P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140698P.  
PR 29-JUN-1999; 99US-0140823P.  
PR 30-JUN-1999; 99US-0140991P.  
PR 01-JUL-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149923P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151068P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158223P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.

Tue Jul 6 09:49:52 2004

us-10-051-902a-20.rag

PR	25-OCT-1999;	99US-0161404P.	PR	23-MAR-1999;	99US-0135788P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-MAR-1999;	99US-0136264P.
PR	25-OCT-1999;	99US-0161406P.	PR	29-MAR-1999;	99US-0136785P.
PR	26-OCT-1999;	99US-0161359P.	PR	01-APR-1999;	99US-0127462P.
PR	26-OCT-1999;	99US-0161360P.	PR	06-APR-1999;	99US-0128234P.
PR	26-OCT-1999;	99US-0161361P.	PR	08-APR-1999;	99US-0128714P.
PR	28-OCT-1999;	99US-0161920P.	PR	16-APR-1999;	99US-0129845P.
PR	28-OCT-1999;	99US-0161922P.	PR	19-APR-1999;	99US-0130077P.
PR	28-OCT-1999;	99US-0161923P.	PR	21-APR-1999;	99US-0130449P.
PR	29-OCT-1999;	99US-0162142P.	PR	23-APR-1999;	99US-0130510P.
Query Match 51.4%; Score 1316; DB 3; Length 491;					
Best Local Similarity 55.6%; Pred. No. 1.e-123;					
Matches 264; Conservative 83; Mismatches 112; Indels 16; Gaps 4;					
QY	19	KYASICAILASMASVILGYDVGMSGAAWYIKKDLNITDVQLEILIGILSLYSLFGSFAG 78	PR	30-APR-1999;	99US-0132407P.
DB	3	KFAGCAIVASIIISIFGYDTGVMSGAQIFIRDDKINDTQIEVLGILNLCALVSLTA 62	PR	04-MAY-1999;	99US-0132484P.
QY	79	ARTSDRIGRLTVFAAVIFPVGSLMCFANVYCMWAGRFVAGVGVGGMIAPVYTAE 138	PR	05-MAY-1999;	99US-0132485P.
DB	63	GKTSDVIGRRYTIASVAVIFLVGSLMGYGNYPVLMVGRCIAGVGVPALMIAPVYSAE 122	PR	06-MAY-1999;	99US-0132486P.
QY	139	ISPAASRGFLTTFPEVINIGILGYLSNFAFARLPLHLGRVWMLAIGAVPSGLLALLVF 198	PR	07-MAY-1999;	99US-0132863P.
DB	123	ISSASHRGFLSLPELCLISGLILGYVSNYCFGKLTLLGWRLMLGIAAFPSLILAFGIT 182	PR	11-MAY-1999;	99US-0134256P.
QY	199	CMPSPRWLVTKGLADARAVLEKTSATPEBAERLADIKAAAGIPKGLDGDVVYVPGKE 258	PR	14-MAY-1999;	99US-0134218P.
DB	183	RMPSERWLVNQGRLAEAKINLVSNTEEEAEFRDILTAAEV-----DVTEI--KE 234	PR	14-MAY-1999;	99US-0134219P.
QY	259	QGGSELO-----VWKKILSPTAVRILLISAVGLHFFQOAGSDSVVOYSARLFKSAG 312	PR	14-MAY-1999;	99US-0134221P.
DB	235	VGGVKKQKHGSKVNRVRLVPRRAVRLILIAAGVHFFEHATGIEAVLYSPRIFFKAG 294	PR	18-MAY-1999;	99US-0134768P.
QY	313	ITDKNKLGVTCAVGVTKFTFFILVATFLLDRAGRRPILLISTGGMIVSLICLGSGLTVAG 372	PR	19-MAY-1999;	99US-0134941P.
DB	295	VVSKDKLLLATVGVGLTKRAFTIIATFLLDKVGRKLLLTSTGGMVFALTSLAVSLTMVQ 354	PR	20-MAY-1999;	99US-0135124P.
QY	373	HHPTKVAWALCIATSLSVIAPFSIGLPGITGVYTSIEIPLOVRALGRAVGVAASNRVT 432	PR	21-MAY-1999;	99US-0135353P.
DB	355	RF--GRLAWSLSLVSYAVAFVSIGLPGITVWYSIEIPFLRAGGASIGVAVNRIM 412	PR	24-MAY-1999;	99US-0135629P.
QY	433	SAVISMTELSLSKAITIGSGSFLLXGIAAFAVFFFTCLPRTGRTLEEMKLFQ 487	PR	25-MAY-1999;	99US-0136021P.
DB	413	NATVSMFLSKAITTGTVGVFFVAGIAAFAVFFFTCLPRTGRTLEEMKLFQ 467	PR	27-MAY-1999;	99US-0136392P.
RESULT 13					
ID	AAG32071 standard; protein; 508 AA.				
XX	AAG32071;				
AC	AAG32071;				
XX	17-OCT-2000 (first entry)				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 38622.				
KW	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
OS	Arabidopsis thaliana.				
XX	EP1033405-A2.				
PN	06-SEP-2000.				
PD	25-FEB-2000; 2000EP-00301439.				
XX	25-FEB-1999; 99US-0121825P.				
XX	05-MAR-1999; 99US-0123180P.				
PR	09-MAR-1999; 99US-0123548P.				

```
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144332P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 03-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147935P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148585P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150894P.
PR 27-AUG-1999; 99US-0151055P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.

PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.4%; Score 1316; DB 3; Length 508;
Best Local Similarity 55.6%; Pred. No. 1.2e-123;
Matches 264; Conservative 83; Mismatches 112; Indels 16; Gaps 4;

QY 19 KYASICAILASMASVILGYDIGWMSGAAMYIKKOLNITDVOLEILIGLISLSLFGSFAG 78
Db 20 KFAFGCAIVASIIIFGYDTGWMGSAQIFIRDDLKINDTQIEVLGAILNLCALVSLTA 79
QY 79 ARTSDRIGRRLTVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGMIAPVYTAE 138
Db 80 GKTSVIGRRYTIASAVIFLVGSYLMGYPNYPVLMVGRGCIAGVGVGFALMIAFVYSAE 139
QY 139 ISPAASRGFLTTFFPEVFINIGILLCYLSNFAPARLPLHLGNRVMLAIGAVPSGLLALV 198
Db 140 ISSASHRGFLSLPELCISGLILGVSNYCFGKLTLLGNWLMGLCIAFSLIILAFIGIT 199
QY 199 CMPESPRMLVLKGRADARAVLEKTSATPEBAERLADIKAAAGIPKGLDGDVWTVPOKB 258
Db 200 RMPESPRMLVMQGRLEAKIMVLVSNTEBAERFRDILTAABV-----DVTEI--KE 251
QY 259 QGGGELQ-----VMKKLILSPTPAVRILLISAVCLHFFQOASGSDSVVQYSARLPK 312
Db 252 VGGGVKKXNHGKSVWRELIVKPRPAVRLLIAAVGIHFEHATGIEAVLVSPRIFKAG 311
QY 313 ITDDNKLGLVTCAGVTKTFFILNATFLLDRAGRPLLLISTGGMIVSLICIGSLTVAG 372
Db 312 VVSKDKLLAIVGVGLTKAFIIATFLLDKVGREKULLTSTGGMVFLTSLAVSLTWVQ 371
QY 373 HHPDTKVAMAVALCIATSLSVIAFSGICLBITGVYVSEIPLQVRALGPVAVSNRVT 432
Db 372 RF--ORLAWALSLSIVSTYAFVFFSIGLGPITWYVSEIPLRLAQAGSAGVAVNRM 429
QY 433 SAVISMFTLSLSKAITIGGSPFLYSGIAAVAWVFFFTCLPRTGRTEEMCKLFG 487
Db 430 NATVSNFSLMTKAITTGGVFFVAGIAVAANWVFFFMPLPETKGLPLEMEKLF 484

RESULT 14
AAG32073
ID AAG32073 standard; protein; 466 AA.
XX
```

us-10-051-902a-20.rag

Tue Jul 6 09:49:52 2004

AC	AA32073;	PR	18-JUN-1999;	99US-0139463P.
XX		PR	18-JUN-1999;	99US-0139750P.
DT	17-OCT-2000 (first entry)	PR	18-JUN-1999;	99US-0139763P.
XX		PR	21-JUN-1999;	99US-0139817P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 38624.	PR	22-JUN-1999;	99US-0139899P.
XX		PR	23-JUN-1999;	99US-0140353P.
XX		PR	23-JUN-1999;	99US-0140354P.
KW	Protein identification, signal transduction pathway; metabolic pathway;	PR	24-JUN-1999;	99US-0140695P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR	24-JUN-1999;	99US-0140823P.
KW	termination sequence.	PR	28-JUN-1999;	99US-0140911P.
XX		PR	29-JUN-1999;	99US-0141287P.
OS	Arabidopsis thaliana.	PR	30-JUN-1999;	99US-0141842P.
XX		PR	01-JUL-1999;	99US-0142154P.
XX		PR	01-JUL-1999;	99US-0142055P.
PN	EP1033405-A2.	PR	02-JUL-1999;	99US-0142390P.
XX		PR	06-JUL-1999;	99US-0142803P.
PD	06-SEP-2000.	PR	08-JUL-1999;	99US-0142920P.
XX		PR	09-JUL-1999;	99US-0142977P.
PF	25-FEB-2000; 2000EP-00301439.	PR	12-JUL-1999;	99US-0143542P.
XX		PR	13-JUL-1999;	99US-0143624P.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0144008P.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144085P.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144325P.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144334P.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144335P.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144352P.
PR	16-APR-1999;	PR	20-JUL-1999;	99US-0144884P.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144814P.
PR	21-APR-1999;	PR	21-JUL-1999;	99US-0145088P.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0145088P.
PR	28-APR-1999;	PR	22-JUL-1999;	99US-0145085P.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145087P.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145089P.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145192P.
PR	05-MAY-1999;	PR	23-JUL-1999;	99US-0145145P.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145218P.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145224P.
PR	07-MAY-1999;	PR	26-JUL-1999;	99US-0145276P.
PR	11-MAY-1999;	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145919P.
PR	14-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	14-MAY-1999;	PR	02-AUG-1999;	99US-0146386P.
PR	18-MAY-1999;	PR	02-AUG-1999;	99US-0146388P.
PR	19-MAY-1999;	PR	02-AUG-1999;	99US-0146389P.
PR	20-MAY-1999;	PR	03-AUG-1999;	99US-0147038P.
PR	21-MAY-1999;	PR	03-AUG-1999;	99US-0147204P.
PR	24-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	25-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	27-MAY-1999;	PR	05-AUG-1999;	99US-0147260P.
PR	28-MAY-1999;	PR	05-AUG-1999;	99US-0147303P.
PR	01-JUN-1999;	PR	06-AUG-1999;	99US-0147416P.
PR	03-JUN-1999;	PR	06-AUG-1999;	99US-0147493P.
PR	04-JUN-1999;	PR	09-AUG-1999;	99US-0147935P.
PR	07-JUN-1999;	PR	09-AUG-1999;	99US-0148171P.
PR	08-JUN-1999;	PR	10-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	PR	11-AUG-1999;	99US-0148319P.
PR	10-JUN-1999;	PR	12-AUG-1999;	99US-0148341P.
PR	14-JUN-1999;	PR	13-AUG-1999;	99US-0148565P.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148684P.
PR	16-JUN-1999;	PR	16-AUG-1999;	99US-0149368P.
PR	17-JUN-1999;	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	PR		



```
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 23-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 08-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 23-OCT-1999; 99US-0162142P.

Query Match 48.7%; Score 1245; DB 3; Length 466;
Best Local Similarity 55.3%; Pred. No. 1.6e-116;
Matches 250; Conservative 79; Mismatches 107; Indels 16; Gaps 4;

Qy 42 MSGAAMYIKKDLNITDVQLEILIGISLVSFSGFAGARTSDRIGRLTVVFAAVIFFVG 101
Db 1 MSGAQIFRDLKINDQIEVLGAILNLCALVSLTAGKTSVDVIGRRYIALSAVIFLVG 60

Qy 102 SLIMGFVAVNCLMAGFVAGVGVGGMIAPIVTAIEISPAASRGILTTPEVFINIGIL 161
Db 61 SVLMGYGNFVPLVMVGCIAGVGVGFALMIAPIVTAIEISSAHRGFLTSLFELCISLIGIL 120

Qy 162 LGYLSNFAFARLPLHLGWRVNLMAIGAVPSGLLALLVFCMPESPRWLVLKRLADARAVILE 221
Db 121 LGYVSNYCFGKLTLLKWLRLMLGIAAPFSLILAFGITRMPESPRVLVMQGRLEAKIMV 180

Qy 222 KTSATPEAAERLADIKAAGIPKGLGDVVTVPGKEQGGGELQ-----VWKKLILSTP 275
Db 181 LVNSTEEBAERFDILTAAEV-----DVTEI--KEVGGGVKKNGKSWFELVIKPR 232
```

```
Qy 276 PAVREILLAVGLHFFQOAGSDSVVOYSARLFSKAGITTDONKLLGVTCAVGVTKTFIL 335
Db 233 FAVRLILAAVGIHFFEHATGIEAVLVYSPRIFKAGVVSKDLLLATVGLGKAFPII 292
Qy 336 VATFLDDRAGRRPLLITSTGMYSLICLSGLTVAGHHPTKVAMAVALCIASLTSLVIA 395
Db 293 IATFLDKVGRKLLTSTGMYFALTSLAVSLTWVQRF--GRLAWALSLSIVSTYAFVA 350
Qy 396 PFSICLGPITGVYTSIEIPLOVRALGFVAGVNASRVTSAVISMFLSLSKAITTGGSPFL 455
Db 351 PFSICLGPITVYTSIEIPFLRLRAQGASIGVAVNRIMNATVSNFSLSKAITTGGVFFV 410
Qy 456 YSGIAAFAVWFFFTCLPSTRGRTLEEMGKLFG 487
Db 411 FAGIAVAFAWFFFTCLPSTRGRTLEEMGKLFG 442

RESULT 15
ADA48320
ID ADA48320 standard; protein; 333 AA.
AC ADA48320;
XX
XX ADA48320;
DT 20-NOV-2003 (first entry)
DE Rice protein conferring disease resistance in plants.
KW disease resistance; pathogen tolerance; plant pathogen; plant; rice.
XX Oryza sativa.
OS
XX
XX WO2003000906-A2.
XX
XX 03-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-IB002453.
XX
XX 22-JUN-2001; 2001US-0300112P.
XX 26-SEP-2001; 2001US-0352277P.
XX 22-MAR-2002; 2002US-0366535P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Katagiri F, Kreps J, Provart N, Rickes D, Zhu T;
XX
XX WPI; 2003-184052/18.
XX
XX N-PSDB; ADA48319.
XX
XX New polynucleotide comprising a plant nucleotide sequence having an open
PT reading frame that encodes a polypeptide associated with disease
PT resistance, useful for conferring resistance or tolerance to a plant
PT pathogen.
XX
XX Claim 10; SEQ ID NO 390; 299pp; English.
XX
XX The invention relates to a novel isolated polynucleotide comprising a
CC plant nucleotide sequence having an open reading frame that encodes a
CC polypeptide associated with disease resistance or its fragment having
CC substantially the same activity as the full-length polypeptide. The
CC polynucleotide of the invention is useful for conferring resistance or
CC tolerance to a plant pathogen. The present sequence represents a protein
CC conferring disease resistance used in the invention.
XX
XX Sequence 333 AA;
SQ
Query Match 30.2%; Score 772.5; DB 6; Length 333;
Best Local Similarity 36.2%; Pred. No. 5.3e-69;
Matches 174; Conservative 53; Mismatches 84; Indels 169; Gaps 6;

Qy 7 AKAVEPRKGNVKKVASICAILASMASVILGVDIGVMSGAAMYIKKDLNITDVQLEILIGI 66
Db 20 AAASLPRN--KYDFFCAVLASMTX-----RAGRSSA-----EDLGVSDAQIEVLGSA 66
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:33:48 ; Search time 22 seconds  
(without alignments)  
1203.824 Million cell updates/sec

Title: US-10-051-902A-20

Perfect score: 2559

Sequence: 1 MASDELAKAVEPRKKNVYK.....AEEADAAKEKVELPSSK 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2559	100.0	513	4	US-09-291-922-20
2	1906.5	74.5	529	4	US-09-291-922-28
3	1872.5	73.2	510	4	US-09-291-922-22
4	1764.5	69.0	539	4	US-09-291-922-26
5	1542	60.3	523	4	US-09-291-922-24
6	1483.5	58.0	543	4	US-09-291-922-30
7	884.5	26.0	488	4	US-10-162-012-46
8	581.5	22.7	514	4	US-09-489-039A-11902
9	576	22.5	167	4	US-09-291-922-18
10	565	22.1	584	2	US-08-928-692-13
11	565	22.1	584	4	US-09-339-972-13
12	551.5	21.6	517	4	US-09-679-686B-18
13	542	21.2	501	4	US-09-489-039A-11731
14	533	20.8	476	4	US-09-489-039A-11933
15	528	20.6	502	4	US-09-679-686B-2
16	511	20.0	518	4	US-09-679-686B-23
17	505.5	19.8	494	2	US-09-031-392-5
18	505.5	19.8	494	3	US-09-299-549-5
19	505.5	19.8	494	4	US-09-610-417-5
20	502.5	19.6	729	4	US-09-291-922-29
21	495	19.3	519	4	US-09-679-686B-24
22	489.5	19.1	510	4	US-09-679-686B-19
23	484	18.9	493	2	US-09-031-392-10
24	484	18.9	493	3	US-09-299-549-10
25	484	18.9	493	4	US-09-610-417-10
26	482	18.8	514	4	US-09-679-686B-22
27	480.5	18.8	524	2	US-08-928-692-12

28	480.5	18.8	524	4	US-09-339-972-12	Sequence 12, Appl
29	477.5	18.7	737	4	US-09-291-922-8	Sequence 8, Appl
30	476.5	18.6	518	4	US-09-679-686B-16	Sequence 16, Appl
31	468.5	18.3	511	4	US-09-679-686B-12	Sequence 12, Appl
32	466	18.2	521	4	US-09-489-039A-9549	Sequence 9549, Ap
33	459.5	18.0	562	4	US-10-162-012-44	Sequence 44, Appl
34	453	17.7	492	2	US-08-355-844-3	Sequence 3, Appl
35	453	17.7	492	5	PCT-US95-16126-3	Sequence 3, Appl
36	453	17.7	747	4	US-09-291-922-2	Sequence 2, Appl
37	443	17.3	504	4	US-09-679-686B-21	Sequence 21, Appl
38	439.5	17.2	509	2	US-09-031-392-6	Sequence 6, Appl
39	439.5	17.2	509	3	US-09-299-549-6	Sequence 6, Appl
40	439.5	17.2	509	4	US-09-610-417-6	Sequence 6, Appl
41	421.5	16.5	534	2	US-09-031-392-4	Sequence 4, Appl
42	421.5	16.5	534	3	US-09-299-549-4	Sequence 4, Appl
43	421.5	16.5	534	4	US-09-610-417-4	Sequence 4, Appl
44	416	16.3	488	2	US-08-928-692-11	Sequence 11, Appl
45	416	16.3	488	4	US-09-339-972-11	Sequence 11, Appl

#### ALIGNMENTS

RESULT 1

US-09-291-922-20

; Sequence 20, Application US/09291922

; Patent No. 6383776

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/09/291,922

; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO. 20

; LENGTH: 513

; TYPE: PRT

; ORGANISM: Zea mays

US-09-291-922-20

Query Match 100.0%; Score 2559; DB 4; Length 513;

Best Local Similarity 100.0%; Pred. No. 4.5e-255;

Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASDELAKAVEPRKKNVYKASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVL 60

DB 1 MASDELAKAVEPRKKNVYKASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVL 60

QY 61 ELILGILSYLSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

DB 61 ELILGILSYLSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

QY 121 AGVGVGGMIAPIVTVTAISPAASRGFTTTPPEVINIGILLYLSNFAFARPLHLGWR 180

DB 121 AGVGVGGMIAPIVTVTAISPAASRGFTTTPPEVINIGILLYLSNFAFARPLHLGWR 180

QY 181 VMLAIGAVPSGLLALVFCMPESPRLVLKGLADARAVLEKTSATPEAAERLADIKAA 240

DB 181 VMLAIGAVPSGLLALVFCMPESPRLVLKGLADARAVLEKTSATPEAAERLADIKAA 240

QY 241 AGIPKGLDGVVTVPEKQGGELQVWKKLILSPTPAVRILLSAVGLHFFQASGSDV 300

DB 241 AGIPKGLDGVVTVPEKQGGELQVWKKLILSPTPAVRILLSAVGLHFFQASGSDV 300

QY 301 VOYSARLPKSGITDDNKLGVTCAGVTKTFFILVATFLDRAGRRRLILISITGMIYS 360

DB 301 VOYSARLPKSGITDDNKLGVTCAGVTKTFFILVATFLDRAGRRRLILISITGMIYS 360

361 L I C L G S G L T V A G H P D T K V A W A V A L C I A S T L S Y I A F F S I G L G P I T G Y T T S E I F P L Q V R A L 420  
D B  
361 L I C L G S G L T V A G H P D T K V A W A V A L C I A S T L S Y I A F F S I G L G P I T G Y T T S E I F P L Q V R A L 420  
Q Y 421 G F A V G V A S N R V T S A V I S M T F L S K A I T I G S F P L Y S G I A A V A W P P T C L P E T R G T L E 480  
D B 421 G F A V G V A S N R V T S A V I S M T F L S K A I T I G S F P L Y S G I A A V A W P P T C L P E T R G T L E 480  
Q Y 481 E M G K L F G M P D T G M A E A E A D A A K E K V V E L P S S K 513  
D B 481 E M G K L F G M P D T G M A E A E A D A A K E K V V E L P S S K 513  
RESULT 2  
US-09-291-922-28  
; Sequence 28, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291.922  
; EARLIER FILING DATE: 1999-04-14  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 28  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-291-922-28  
Query Match 74.5%; Score 1906.5; DB 4; Length 529;  
Best Local Similarity 73.6%; Pred. No. 9,3e-188;  
Matches 373; Conservative 58; Mismatches 71; Indels 5; Gaps 3;

Q Y 1 M A S D E L A K - A V E P R K K N V K Y A S I C A I L A S M A S V I L G Y D I G V M S G A M Y I K K D L N I T D V 58  
D B 19 M A S A A L P E A V P R K N K N F Y A F T C A L C A S M A T I V L G Y D V G V M S G A S L Y I K R D L Q I T D V 78  
Q Y 59 Q L E I L I G I L S L Y S L F G S F A G A R T S D R I G R R L T V V F A A V I F F V G S L L M G F A V N Y M L M A G R 118  
D B 79 Q L E I M M G I L S V A L I G S F E L G A R T S D W V G R R V T V V F A A I F N N G S L L M G F A V N Y M L M V G R 138  
Q Y 119 F V A G V G Y G G M I A P V T Y A E I S P A A S R G L T T F P E V F I N I G I L L Y S N F A F A R L P L H L G 178  
D B 139 F V T G I G V Y A I M A V P V T Y P V S P A S A R G F L T S F T E V F I N V G I L L Y S N V A F A R L P L H L S 198  
Q Y 179 W R V M A I G A V P S G L L A L V F C M P E S P R W L V I K G L A D A R A V L E K T S A T P E A A E R L A D I K 238  
D B 199 W R V M I G I A V P S A L L A M V F G M P E S P R W L V M K G L A D A R A V L A K T S D T P E A V E R L D O I K 258  
Q Y 239 A A A G I P K G L D G D V V T P K E G G E L Q W K K L I L S P T P A V R R I L L S A V G L H F F Q A S G S D 298  
D B 259 A A A G I P R E L D G D V V M P - K T K G G E K Q W K E L I F S P T P A M R I L L A A L G I H F F Q A T G S D 317  
Q Y 299 S V Q V S A R L F K S A G I T D D N K L I L G V T C A V G V T K T F F I L V A T F L D R A G R R P L L I S T G M I 358  
D B 318 S V L Y S P R V F S A G I T G N H L I G A T C A M G V M K T L F I L V A T F Q L D R V G R R P L L L S T A G M L 377  
Q Y 359 V S L I G S G L T V A G H P D T K V A W A V A L C I A S T L S Y I A F F S I G L G P I T G Y T T S E I F P L Q V R 418  
D B 378 A C L I G L G T L V G R H P D A K V P F A I G L C I V I L A V S F S I G L G F L S V T S E V F P L R V R 437  
Q Y 419 A L G F A V G V A S N R V T S A V I S M T F L S K A I T I G S F P L Y S G I A A V A W P P T C L P E T R G T 478  
D B 438 A L G F A L G T S C N R V T S A A V S M G S F L S K A I T I G S F P L Y S G I A A V A W P P T C L P E T R G T L P 497

Q Y 479 L E M G K L F G M P D T G M A E A E A D A A K E K 505  
D B 498 L E E I G K L F G M T D T - A V E A Q D T A T K D K 522  
RESULT 3  
US-09-291-922-22  
; Sequence 22, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291.922  
; EARLIER FILING DATE: 1999-04-14  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 22  
; LENGTH: 510  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; NAME/KEY: UNSURE  
; LOCATION: (102)  
US-09-291-922-22  
Query Match 73.2%; Score 1872.5; DB 4; Length 510;  
Best Local Similarity 72.7%; Pred. No. 2.8e-184;  
Matches 372; Conservative 60; Mismatches 77; Indels 3; Gaps 2;

Q Y 1 M A S D E L A K A V E P R K K N V K Y A S I C A I L A S M A S V I L G Y D I G V M S G A M Y I K K D L N I T D V Q L 60  
D B 1 M A S A A L P E A V P R K K N V R F A F A C A I L A S M T S I L L G Y D I G V M S G A S L Y I K K D F N I S D G K V 60  
Q Y 61 E I L I G I L S L Y S L F G S F A G A R T S D R I G R R L T V V F A A V I F F V G S L L M G F A V N Y M L M A G R F V 120  
D B 61 E V L M G I L N L Y S L I G S F A A G R T S D W I G R R V T V F A A V I F F A G X F L M G F A V N Y M L M F G R F V 120  
Q Y 121 A G V G V G Y G M I A P V T Y A E I S P A A S R G F L T T P P E V F I N I G I L L Y S N F A F A R L P L H L G W R 180  
D B 121 A G I G V G Y A L M I A P V T Y A E V S P A S A R G F L T S P P E V F I N F I L L G Y V S N V A F S R L P L N L G W R 180  
Q Y 181 V M A I G A V P S G L L A L V F C M P E S P R W L V I K G L A D A R A V L E K T S A T P E A A E R L A D I K A A 240  
D B 181 I M L G I A A P S V L L A M V L G M P E S P R W L V M K G L A D A K V L E K T S D T A E A E R L A D I K A A 240  
Q Y 241 A G I P K G L D G D V V T P K E G G E L Q W K K L I L S P T P A V R R I L L S A V G L H F F Q A S G S D S V 300  
D B 241 A G I P E L D G D V V T P - K R G S G N E K R V K E L I L S P T P A M R I L L S G I G I H F F Q H A L G I H S V 299  
Q Y 301 V Q Y S A R L F K S A G I T D D N K L I L G V T C A V G V T K T F F I L V A T F L D R A G R R P L L I S T G M I V S 360  
D B 300 V F Y S P A V F K S P G L T N D X H F L G T T W F F G V T K R L F I L A I F F I D G V G R R P L L G S T G G I L S 359  
Q Y 361 L I C L G S G L T V A G H P D T K V A W A V A L C I A S T L S Y I A F F S I G L G P I T G Y T T S E I F P L Q V R A L 420  
D B 360 L I C L G A G L I V V G H P D A K I P W A I G L S I A T L A V A F F S I G L G P I T W Y S S E I F P L Q V R A L 419  
Q Y 421 G F A V G V A S N R V T S A V I S M T F L S K A I T I G S F P L Y S G I A A V A W P P T C L P E T R G T L E 480  
D B 420 G C S L G V A A N R V T S G V I S M T F L S K A I T I G S F P L Y S G I A A V A W P P T C L P E T R G T L E 479  
Q Y 481 E M G K L F G M P D T G M A E A E A D A A K E K V V E L P S S 512  
D B 480 E M S K L F G - D T A A S E S D E P A K E K K V E M A A T 509

RESULT 4

US-09-291-922-26  
; Sequence 26, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 26  
; LENGTH: 539  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-291-922-26

Query Match 69.0%; Score 1764.5; DB 4; Length 539;  
Best Local Similarity 68.8%; Pred. No. 4.2e-173;  
Matches 351; Conservative 63; Mismatches 91; Indels 5; Gaps 2;

QY 6 LAKAVEPRKKNVYASICAILASMASVILGYDVGWMSGAAMYIKDLNITDVOLEILIG 65  
DB 31 LPAAVEPRKKNVYASICAILASMTSILLGYDVGWMSGASLYIQDKINDTQLEVLNG 90  
QY 66 ILSYLSFGSPAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLAGREAVAGVY 125  
DB 91 ILNVYLSITGSPAAGRTSDWIGRRFTVFAAVIFFAGALIMGFSVNYAMLFGREAVAGVY 150  
QY 126 GYGGMIAPVYTAETSPASRGFLTFEVRNIGILLGYLSNFAFARLPLHLGHRVLMIAI 185  
DB 151 GYALMIAPVNTGEVSPASRGVLTSPFEVFNFGILLGYVSNFAFARLSRLGHRVLMIGI 210  
QY 186 GAVPSGLLALVFCMPSPRVLVKGRLADARAVLEKTSATPEBAERLADIKAAAGIPK 245  
DB 211 GAVPSVLLAFVNLGVPSPRVLVKGRLADAKVLAKTSDTPEBAERLADIKTAAGIPL 270  
QY 246 GLDGDVTVPKGEGGELQWKKLLISPTPAVRILLSAVGLHFFQOAGSDSVVOYSA 305  
DB 271 GLDGDVFPVKNKGSSEKRYLKDILSPTTAMRHILLIAGTGHFFQOAGSDVAVLYSP 330  
QY 306 RLFSAGITDDNKLGVTCVAGVTKTFFILVATLDRAGRRPLLISGTGMIVSLICIG 365  
DB 331 LVFSGAGITGSRLRGTVAGVANTVPIVATLDRIRRRPLLISGTGMIVSLVGLA 390  
QY 366 SGLTVAGHHPDKVAVAVLCIASTLSYIAFFSIGLGPITGVYTSIEFPLQVRALGPAVG 425  
DB 391 TGLTVISRHPDEKITWAVLCIFCIMAAYVAFSIGLGPITWVYSSEIFPLHVRALGCSLG 450  
QY 426 VASNRVTSVAVSMFTLSLSKAITTIGGSFFLYSGIAAVAVVFFFTCLPSTRGRITLSENGKL 485  
DB 451 VAVNRVTSVAVSMFTLSLSKAITTIGGSFFLYSGIAAVAVVFFFTCLPSTRGRITLSENGKL 510  
QY 486 FGMPDPT---GMAEEAEADAEEKVVELPSS 512  
DB 511 FGNTATHKQGAEEADDDAG--EKKVEMAAT 538

RESULT 5  
US-09-291-922-24  
; Sequence 24, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins

FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 24  
; LENGTH: 523  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-291-922-24

Query Match 60.3%; Score 1542; DB 4; Length 523;  
Best Local Similarity 62.0%; Pred. No. 3.6e-150;  
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKNVYASICAILASMASVILGYDVGWMSGAAMYIK 50  
DB 1 MTEGKLVEAAEAHKTLDQDPFKKRNKYAFACAMLASMTSILLGYDVGWMSGAAMYIK 60  
QY 51 KDLNITDVOLEILIGILSYLSFGSPAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVN 110  
DB 61 RDLKVSDEQTEIILLGILINLYSLIGSLAGRTSDWIGPRVTIVFAGTIFVFGALLMGFSPN 120  
QY 111 YGMLAGRFVAGVGVGGMIAPVYTAETSPASRGFLTFEVRNIGILLGYLSNFAF 170  
DB 121 YSFLMFRFVAGIGIGYALMIAPVYTAETSPASRGFLTFEVRNIGILLGYLSNFAF 180  
QY 171 ARPLHLGHRVLMIAI GAVPSGLLALVFCMPSPRVLVKGRLADARAVLEKTSATPEEA 230  
DB 181 SKLTAKVGRMMLGVAIPSVLLTVGVLAMPSPRVLVKGRLGKARKVNLKNTSDSKEEA 240  
QY 231 AERLADIKAAAGIPKGLDGDVTVPKGEGGELQWKKLLISPTPAVRILLSAVGLHF 290  
DB 241 QURLASIKQAAGIPESCDNDVQVKNQSGEG---VWKEFLYPTPAIRHIVIAALGIFH 297  
QY 291 FQOAGSDSVVOYSARLFSAGITDDNKLGVTCVAGVTKTFFILVATLDRAGRRPLL 350  
DB 298 FQOAGSDVAVLYSPRIFEKAGITNDTHKLATVAGFVKTVFLLAATFLDRVGERPLL 357  
QY 351 LISTGMIVSLICLSGLTVAGHHPDKVAVAVLCIASTLSYIAFFSIGLGPITGVYTS 410  
DB 358 LSSVGMVLSLTLTALSLTVI-DHSEKRLMWAVGSSIAMVLAYVATFSGAGPITWVYSS 416  
QY 411 EIFFLOVRALGPAVGAVASNRVTSVAVSMFTLSLSKAITTIGGSFFLYSGIAAVAVVFFFTC 470  
DB 417 EIFFLRLRQGAAGAVAVNRVTSVAVSMFTLSLRAITIGGAFFLYCGIATVGVIFFYIV 476  
QY 471 LPETGRITLSENGKLF 487  
DB 477 LPETGRITLSENGSFG 493

RESULT 6  
US-09-291-922-30  
; Sequence 30, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 30  
; LENGTH: 549

```

; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-291-922-30

Query Match      58.0%; Score 1483.5; DB 4; Length 549;
Best Local Similarity 61.6%; Pred. No. 4.2e-144; Indels 11; Gaps 5;
Matches 299; Conservative 63; Mismatches 112;

; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-162-012-46

Query Match      26.0%; Score 664.5; DB 4; Length 488;
Best Local Similarity 34.9%; Pred. No. 1e-59;
Matches 177; Conservative 91; Mismatches 174; Indels 65; Gaps 15;

QY    25 AILASM-ASVILGYDVGMSG-----AAMVKDLNLTVDQLILIGILS 68
DB    2 ALVAALGGFLFYDVTGVIGFLALIDFLPRFGLTSSGALAELVGYSTVLTLGLWVSIF 61
QY    59 LYSLPGSFAGARTSDRIGRRLLTVFAAVIPFGVSLMGFAVNVG-----MLMAGRFF 120
DB    62 LGRLIGSLFAGKDRFCGRKKSLIALVLFIHALLSGAAPGYTTICLMWFYLLIVGRVL 121
QY    121 AGVGVGYGMIAPVVYTABISPAASRGFLTTPPEVFINIGILLGYLSNFAPARPLH---- 176
DB    122 VGLGVGGASVLPMVYISEIAPKALRGALGSYLQAITIGILVA-----AIIGLGNKTNN 176
QY    177 -----LGWRVWLGAIGVPSCGLLALLVFCMPESPRMLVXGRADARAVLEKTSATPEEA 230
DB    177 DSALNSWGRIPLGLQLPALLLIIGLFLUPESPRMLVEKGKEAEAREVIAKURGV-EDV 235
QY    231 ASRLADIKAAAGIPKGLDGDDVTVTPGKEQGGEIQVWKKLILSPT--PAVRRIILLSAVGLH 289
DB    236 DQEIEIKAE-----LEATV-----SEEKAGKAS--NGELFRGTRFKVQRLLMGVMQL 283
QY    290 FFQASGSDSVQVVSARLFKAGITTDKNLLGVTCANGVTKTFPIIVA--TELDLRAGRRP 348
DB    284 AFQQLTGAINAFYXSPTTFKSVGSDSVASLLVTVIIIVGWNVFTFFVALIFLVDFRGRRP 343
QY    349 ILLISTGMIVSLICLGSGLTVA-----GHHPDTKVAWAVALCIASTLVIATFFSIGLG 402
DB    344 LULLGAAGAHCFLILGASIGVALLLNKPDKPSKAAGIVA--IVFILLFIAFFALGWG 401
QY    403 PITGVYVTEIFLPQVRALGFVAGVSNRVTSVAVISMFTLSLKAI--TIGGSFPL-YSGI 459
DB    402 PIPWVILSELFPYKRSKALATALAANWLANFIIGFLFPVITGAIGLGGYVFLVPAFL 461
QY    460 AAVALVFEFTCLPETRGRTLEEMKLF 486
DB    462 LVLFILFVFFVPEYKGTLEIELF 488

RESULT 7
US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05

```

## RESULT 8

US-09-489-039A-11902  
; Sequence 11902, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:

; APPLICANT: Gary Bretton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489.039A

; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 11902

; LENGTH: 514

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-11902

Query Match 22.7%; Score 581.5; DB 4; Length 514;

Best Local Similarity 29.9%; Pred. No. 3.9e-51;

Matches 147; Conservative 101; Mismatches 200; Indels 43; Gaps 9;

Qy 1 MASDELAKAVEPRKGNVKNYAS-ICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQ 59

Db 42 MTSISNDSTLSPRTQDTRRMWVFSIAAAVAGLLFGLDIDGIVSGALPFTIDHFTLSSQL 101

Qy 60 LEILIGILSYLSLFGSFAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRF 119

Db 102 QEWVSSNMGLAAALFNGWLSFRLGRKYSLMAGAVLFWAGSISGSAASVEVLLVARV 161

Qy 120 VAGVGVGGMIAPVYTAETISPAASRGFLTTFPEVFNIGILGYSNFAFARLPLHLG- 178

Db 162 VLGAVGASVTAFLYLSMASENVKGMISMYQLMVTGLVLAFLSDTAFS-----YSGN 217

Qy 179 WRVMLATGAVPSGLLALLVFCMPSPRWLVKGLADARAVLEKTSATPEAAERLADIK 238

Db 218 WRAMLGLVALPAVILILVFLPNSPRWLAEGKSHIEAEVLRMLRDTSEKARDELNEIR 277

Qy 239 AAAGIPKGLDGVTVTPGKEGGGELQVWKLLISPTPAVRRLILLSAVGHFFQAGSGD 298

Db 278 ESKL-----KQGG-----WALFKINRN--VRRVFLGMLLOAQVQFTGMN 316

Qy 299 SVVOYARLFKSGITDDNKLGLVTCVAVGVTKTFILVATFLDLDRAGRRLPLLIISGMI 358

Db 317 IIMYAPRIFKQWAGFTTTEQQMIATLVVGLTFEPATFIAVFTVDKAGRKPAKIGFSVMA 376

Qy 359 VSLICLG-----SGLTVAGHEDTKVANAVALCIASITLSYIAFFSIGLGPITGVYSE 411

Db 377 LGTLVLGVLQDFNGTASSG-----LSW---LSVGMWMCAGYAMSAAPVWVILCSE 427

Qy 412 IFPIQVRALGFVAGVAGSNRVTSAVISMTFLSLKAITIGGSFFLYSGIAAVAVWVFFFTCL 471

Db 428 IQPLKCRDFGJTCSTTINWVSNMIIIGATFTLLDAIGAAGTFWLYTALNAVFIGIFWLI 487

Qy 472 PETGRILEEM 482

Db 488 PETKNVTLLEHI 498

## RESULT 9

US-09-231-922-18

; Sequence 18, Application US/09291922

; Patent No. 6383776

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/09/291,922

; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 18

; LENGTH: 167

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (34)

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (85)

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (98)

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (112)

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (151)

US-09-291-922-18

Query Match

Best Local Similarity 22.5%; Score 576; DB 4; Length 167;

Matches 114; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

Qy 1 MASDELAKAVEPRKGNVKNYAS-ICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQ 60

Db 10 MASAPLPAIEPGKGNVKFAFACXILASMTSILLGYDIGVMSGASLYIKKDLKISDVKL 69

Qy 61 EILIGILSYLSLFGSFAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

Db 70 EILMGILNYSLTGSXAAGRTSDWIGRRXTIVFAAVIFFAGAXLMGFAVNYWMLMGRFV 129

Qy 121 AGVGVGGMIAPVYTAETISPAASRGFLTTFPEVFI 156

Db 130 AGIGVGYALMIATVYTAEVSPXASRGFLTSPEVFI 165

## RESULT 10

US-08-928-692-13

; Sequence 13, Application US/08928692

; Patent No. 5958727

; GENERAL INFORMATION:

; APPLICANT: Brody, Howard

; APPLICANT: Yaver, Deborah S.

; APPLICANT: Lamsa, Michael

; APPLICANT: Hansen, Kim

; TITLE OF INVENTION: Methods for Modifying the Production of

; NUMBER OF SEQUENCES: 80

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5958727o No. 5958727disk of No. 5958727th America, Inc.

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: Fast-SEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,692

; FILING DATE: 12-SEPT-1997

; CLASSIFICATION: 435

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5958727e
; US-08-928-692-13

Query Match 22.1%; Score 565; DB 2; Length 584;
Best Local Similarity 29.5%; Pred. No. 2.4e-49;
Matches 150; Conservative 109; Mismatches 183; Indels 66; Gaps 13;

QY 27 LASMASVILGYDIGVMSGAAMYIKDLN---ITDVQLEILIGILSLYSFSGFAGARTSD 83
Db 92 VASISGFMFGYDTGYISSALISIGTDLHKVLTGKEKEIVTAATSLGALITSIIFAGTAAD 151
QY 84 RIGRLTVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGYGMIAPIVYTAISPA 143
Db 152 IFGRKRLGMSNLMFVIGAILQVSAHTFWQAVGRLINGFGVGLSLIAPLFISEIAPKM 211
QY 144 SRGFLTTPPEVINIGIL-----LQYLSNFAFARLPLHLGWRVMAIGAIVPSGLLAL 195
Db 212 IRGRLTVINSMLTGGQLVAYCGGAGLNYYNN-----GWRILVGLSLIPTAVQFT 261
QY 196 LVFCMPESPRLVLKRLADARAVLEKT-SATPEEAERLADIKAAAGIPKGLDGDVVT- 253
Db 262 CLCFLEDPTRVYVYMGDLARATEVLKRSYTDTSIEIERKVE-----ELVTL 308
QY 254 ---VPGKEGGGELQVKKL-ILSPTPAVRRIILSAVGLHFFQOAGSDSVVQYSAFLFK 309
Db 309 NQSIIPGKNV---PEKVMNTIKELHTVPSNLRALIIGCGLOAIQOFTGWNLSMYSFGTIFE 365
QY 310 SAGITDDNKLGLVTCVAGVTKTFFILVATFLDRAGRRPLLLISTGGMIVSLICLGSGLT 369
Db 366 TVGFKNSS---AVSIIVSGTNFIPLVAFPSIDKIRRTILLIGLPGMTALVV---CS 418
QY 370 VAGHHPDTKVAVAVALCIASLTLS-----YIAFFSIGLGPITGVYTSIFPLQV 417
Db 419 IAFHFLGKFDGAVAVVSSGFSWGIVIIIVFAAFYALGIGTVPW-QQSELPQNV 477
QY 418 RALGFVAGVASNRVTSAVISMTFLSKAITIGSFYLYSGIAAVAVVFFTCPLPETRGR 477
Db 478 RGIGTSYATATNWAGSLVIATFTMLQNTITAGTFAFFAGLSCLSTIFCYFCYFELSGL 537
QY 478 TLEEMGKLFQMPDPTGMAEEADAARKEK 505
Db 538 ELEEVQTIL---KQGFNIKAKALAKR 562

```

## RESULT 11

```

US-09-339-972-13
; Sequence 13, Application US/09339972
; Patent No. 6323002
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6323002o No. 6323002disk of No. 6323002th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York

```

```

; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/339,972
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/928,692
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6323002e
; US-09-339-972-13

Query Match 22.1%; Score 565; DB 4; Length 584;
Best Local Similarity 29.5%; Pred. No. 2.4e-49;
Matches 150; Conservative 109; Mismatches 183; Indels 66; Gaps 13;

```

```

QY 27 LASMASVILGYDIGVMSGAAMYIKDLN---ITDVQLEILIGILSLYSFSGFAGARTSD 83
Db 92 VASISGFMFGYDTGYISSALISIGTDLHKVLTGKEKEIVTAATSLGALITSIIFAGTAAD 151
QY 84 RIGRLTVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGYGMIAPIVYTAISPA 143
Db 152 IFGRKRLGMSNLMFVIGAILQVSAHTFWQAVGRLINGFGVGLSLIAPLFISEIAPKM 211
QY 144 SRGFLTTPPEVINIGIL-----LQYLSNFAFARLPLHLGWRVMAIGAIVPSGLLAL 195
Db 212 IRGRLTVINSMLTGGQLVAYCGGAGLNYYNN-----GWRILVGLSLIPTAVQFT 261
QY 196 LVFCMPESPRLVLKRLADARAVLEKT-SATPEEAERLADIKAAAGIPKGLDGDVVT- 253
Db 262 CLCFLEDPTRVYVYMGDLARATEVLKRSYTDTSIEIERKVE-----ELVTL 308
QY 254 ---VPGKEGGGELQVKKL-ILSPTPAVRRIILSAVGLHFFQOAGSDSVVQYSAFLFK 309
Db 309 NQSIIPGKNV---PEKVMNTIKELHTVPSNLRALIIGCGLOAIQOFTGWNLSMYSFGTIFE 365
QY 310 SAGITDDNKLGLVTCVAGVTKTFFILVATFLDRAGRRPLLLISTGGMIVSLICLGSGLT 369
Db 366 TVGFKNSS---AVSIIVSGTNFIPLVAFPSIDKIRRTILLIGLPGMTALVV---CS 418
QY 370 VAGHHPDTKVAVAVALCIASLTLS-----YIAFFSIGLGPITGVYTSIFPLQV 417
Db 419 IAFHFLGKFDGAVAVVSSGFSWGIVIIIVFAAFYALGIGTVPW-QQSELPQNV 477
QY 418 RALGFVAGVASNRVTSAVISMTFLSKAITIGSFYLYSGIAAVAVVFFTCPLPETRGR 477
Db 478 RGIGTSYATATNWAGSLVIATFTMLQNTITAGTFAFFAGLSCLSTIFCYFCYFELSGL 537
QY 478 TLEEMGKLFQMPDPTGMAEEADAARKEK 505
Db 538 ELEEVQTIL---KQGFNIKAKALAKR 562

```

RESULT 12



US-09-679-686B-18  
; Sequence 18, Application US/09679686B  
; Patent No. 6624343  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen M.  
; APPLICANT: Lightner, Jonathan E.  
; APPLICANT: Rafalski, J. Antoni  
; APPLICANT: Thorpe, Catherine J.  
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS  
; FILE REFERENCE: BB160 US NA  
; CURRENT APPLICATION NUMBER: US/09/679,686B  
; PRIOR FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: 60/081,131  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: PCT/US99/07561  
; PRIOR FILING DATE: 1999-04-07  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 18  
; LENGTH: 517  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-679-686B-18

Query Match 21.6%; Score 551.5; DB 4; Length 517;  
Best Local Similarity 30.2%; Pred. No. 4.9e-48;  
Matches 146; Conservative 90; Mismatches 196; Indels 51; Gaps 10;  
QY 26 ILASMASVILGYDIGVMSGAA-----MYKKDLNITDVQ-----LEILIG 65  
DB VVAATGGILFYDIGISGVTSMNPFLLKFFPEVYDKKMGKSASQYCKVDNQLQFTS 87  
QY 66 ILSYSLFGSPAGATSDRIGRRLTVVFAAIVFVGSLLMGFAVNYGMLMAGRVAGV 125  
DB 88 SLYAALVSSFFAATVTVVGRKMSMTGGTLTFLIGALNGAENIAMLVGRILLGV 147  
QY 126 GYGMIAAPVYTAELSPAASRGFLTTFPEVFINIGILLGYSNFAFARLP LHLGRVNLAI 185  
DB 148 GFANQSPVVLSEMAPARLRGMLNIGQLMTITIGLAAALINYDTNKIKAGYGRISLAI 207  
QY 186 GAVPSGLLALLVFCMPSPRWLVKGRADARAVLEKTSATPEAAERLADIKAAAGIPK 245  
DB 208 AAVFAGITLGSFLLPTPNSLIERGHPPEAARMNRIGSDVDISEYADLVVASE--- 264  
QY 246 GLDGDVTVPGKEGGELQVKKLILSPTPAVRRIILSAVGLHFFQOAGSDSVVOYSA 305  
DB 265 --ESKLVQHP-----WRNLI---QRKYRQPLTMAIMPFFQQLGIVNIMFYAP 308  
QY 306 RLFSAGITDDNKLLGVTCVGVTKTFLLVATFLLDRAGRRLPLLISITGGMIVSLICLG 365  
DB 309 VLFEPLGFKGDASLMSAV-ITGLVNVFATLVSVFTVDRGRKRLFLQGGTQMLLSQLVVG 367  
QY 366 S-----GLTVAGHPDTKVAVALCIASLSVIAFFSICLGPITGVYTSIEIFLPQVRA 419  
DB 368 TLIAVKGTSGVEMPCKGYAA-AVVLPIG---LVVAGFAWSGPLGLWLPSEIFPLEIRP 423  
QY 420 LGFVAVGASNRVTSVAVSMFTLSLKAITIGSFYSLGIAAVALVAVFFFTCLPSTRGTL 479  
DB 424 AGQGINVSNMLFTFVIAQAFLLCHMKF-GLFYFFAGVWVIMTVFIALFLPETKNVPI 482  
QY 480 EEM 482  
DB 483 EEM 485

RESULT 13  
US-09-489-039A-11731  
; Sequence 11731, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11731  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11731

Query Match 21.2%; Score 542; DB 4; Length 501;  
Best Local Similarity 29.5%; Pred. No. 4.5e-47;  
Matches 148; Conservative 95; Mismatches 204; Indels 54; Gaps 12;  
QY 22 SICAILASMASVILGYDIGVMSGAAYIKKDLNITD-VOLEIILIGILSYLSPGSPAGAR 80  
DB 38 TIC-LVNAACGGLFGYDWWVVGAKFFYEANFSITDPAQSGWAMSSALICGIFGALISGW 96  
QY 81 TSDRIGRRLTVVFAAIVFVGSLLMGFAVNYGMLMAGRVAGVGVGGYGGMIAPVYTAELS 140  
DB 97 CADKLGKLPILSAVLFSASAWGTAVASHFDMFVVYRVGVGVIGLASALSPLYIAEVS 156  
QY 141 PAASRGFLTTFPEVFINIGILLGYSNFAFARLP L-----HLGRVNLAI 185  
DB 157 PAEKRGRFVAVNQLTIVIGVLAQLINLMIAE-PVEPGATQOMIVDSWNGQMGWRMFGA 215  
QY 186 GAVPSGLLALLVFCMPSPRWLVKGRADARAVLEKTSATPEAAERLADIKAAAGIPK 245  
DB 216 ELVPALAFVLVMEFFVFPSPRWLMKAGKPERASAALERIGSA--DYADRI-----LREIAH 268  
QY 246 GLDGDVTVPGKEGGELQVKKLILSPTPAVRRIILSAVGLHFFQOAGSDSVVOYSA 305  
DB 269 TLEKD---NNKVSYGALL-----APQVKPIVIGMVLAIPOQWCGINVFNVAQ 314  
QY 306 RLFSAGITDDNKLLGVTCVGVTKTFLLVATFLLDRAGRRLPLLISITGGMIVSLICLG 365  
DB 315 EIFASAGP-DINSTLSIVATGVNVLVFTIALPLVDKIGRKLMLLGNASGLTLYV--- 370  
QY 366 SGLTVAGHPDTKVAVALCIASLSVIAFFSICLGPITGVYTSIEIFLPQVRAFGVAVG 425  
DB 371 --LIAGAVAMGIMGWPLVLLV---LAAIAVALTAPVTWVLLAEIFPNRVRGLAMSLG 424  
QY 426 VASNRVTSVAVSMFTLSLKAITIGSFYSLGIAAVALVAVFFFTCLPSTRGTLBEMGKL 485  
DB 425 TLALMTACFLTYTTPPLNAGLGAAGSFLLYGVI CAAGYLYILRNVPETKGTITLEALSEQ 484  
QY 486 FGMPDTGMAEEAEADAAREKV 506  
DB 485 LAQRHTGV-----NAAKQEQM 500

RESULT 14  
US-09-489-039A-11933  
; Sequence 11933, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11933  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11933

```

Query Match      20.8%; Score 533; DB 4; Length 476;
Best Local Similarity 27.9%; Pred. No. 3.5e-46;
Matches 136; Conservative 99; Mismatches 206; Indels 46; Gaps 9;

QY 13 RKKG--NVKYSACILASMASVILGYDYGMSGANMYIKKDLNITDVQLEILLIGILSLY 70
Db 17 KQGRSNTMTTFVCFALAGLFLGLDGIAGALPIANEFOLSAHTQEWVSSMMFG 76
QY 71 SLFGSFACARTSDRIGRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRFVAGVGVGCGM 130
Db 77 AAVGAVGSGWLSFKLGRKXSLMIGAILFVAGSLFSAAPNVVEILLVSRVLLGLAVGVASY 136
QY 131 IAPVYTAISPAASRGFTTTPPEVFINIGILLGYSNFAFARLPHLG-WRVMLAIGAVP 189
Db 137 TAPLYLSIAPEKIRGSMISVYQLMTITIGILGAYLSDTAFS-----YSGAWRWMLGVIIIP 192
QY 190 SGLLALLVFCMPESPRLVLKGLRADARAVLEKTSATPEEAERLADIKAAAGIPKGLDG 249
Db 193 AVLLIGVIFLDPSPRWFAKRRFVDAERVLLRLRDTSAEAKRELDIRESLKV----- 246
QY 250 DVVTVPKGQGGELQVWKILISPTPAVRRIILSAVLGHFFQOASGSDSVVQYSARLFK 309
Db 247 -----KQSG-----WS--LFKDNSFRRAVFLGILLQVMOQFTGNVINYAPKIFE 291
QY 310 SAGITDDNKLGLVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGMIVSLICLGSGLT 369
Db 292 LAGYANTTEQMWGTVIVGLTNVLTATFIAIGLVDRWGRKPTLILGFIYMAAGNGVLGTMMH 351
QY 370 VAGHPDTKVAVAVALCIASITLSYIAFFSIGLPGITGYTSEIPLQVRALGFVGVASN 429
Db 352 I-GHSST-----AQYIAVLMLLMFIVGFAMSGAGPLIWLVCSEIOLPKGRDFGTCSTATN 406
QY 430 RVTSAVISMTPLSKATIGSGSPFLYSGIAAVAVVFFTCLPETR-----G 476
Db 407 MIANNIVGATFILTMLNSLGSAFTFWVYGGVNLVFIILLTLIPETKRVLSHIERNLMOG 466
QY 477 RTLEEMG 483
Db 467 RPLREIG 473

RESULT 15
US-09-679-686B-2
; Sequence 2, Application US/0967966B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (488)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (493)
; OTHER INFORMATION: Xaa = any amino acid
; US-09-679-686B-2

```

Search completed: June 30, 2004, 18:37:51  
Job time : 36 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:33:08 ; Search time 50 Seconds  
(without alignments)  
2900.556 Million cell updates/sec

Title: US-10-051-902A-20  
Perfect score: 2559  
Sequence: 1 MASDELAKAVEPRKGNVYK.....AAEADRAAEKVVLPSSK 513

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2559	100.0	513	US-10-051-902-20	Sequence 20, Appl
2	2559	100.0	513	US-10-051-909-20	Sequence 20, Appl
3	1906.5	74.5	529	US-10-051-902-28	Sequence 28, Appl
4	1906.5	74.5	529	US-10-051-909-28	Sequence 28, Appl
5	1872.5	73.2	510	US-10-051-902-22	Sequence 22, Appl
6	1872.5	73.2	510	US-10-051-909-22	Sequence 22, Appl
7	1764.5	69.0	539	US-10-051-902-26	Sequence 26, Appl
8	1764.5	69.0	539	US-10-051-909-26	Sequence 26, Appl
9	1626	63.5	327	US-10-051-114-68399	Sequence 68399, A
10	1612.5	63.0	356	US-10-051-114-61926	Sequence 61926, A
11	1597	62.4	380	US-10-051-114-39509	Sequence 39509, A
12	1542	60.3	523	US-10-051-902-24	Sequence 24, Appl
13	1542	60.3	523	US-10-051-909-24	Sequence 24, Appl
14	1487.5	58.1	517	US-10-051-114-56035	Sequence 56035, A
15	1483.5	58.0	548	US-09-774-381-40	Sequence 40, Appl

16	1483.5	58.0	549	13	US-10-051-902-30	Sequence 30, Appl
17	1483.5	58.0	549	13	US-10-051-909-30	Sequence 30, Appl
18	1483.5	57.4	541	12	US-10-051-902-733	Sequence 66733, A
19	1441	56.3	580	12	US-10-051-114-67056	Sequence 67056, A
20	1380	53.9	546	12	US-10-051-114-63789	Sequence 63789, A
21	1293.5	50.8	424	12	US-10-051-999-199875	Sequence 199875, A
22	1293.5	50.1	488	12	US-10-051-999-169603	Sequence 169603, A
23	1279.5	50.0	502	12	US-10-051-114-51712	Sequence 51712, A
24	1195.5	46.7	407	12	US-10-051-114-49353	Sequence 49353, A
25	1193	46.6	531	12	US-10-051-114-182839	Sequence 182839, A
26	1182	46.2	281	12	US-10-051-114-50090	Sequence 50090, A
27	1170	45.7	533	15	US-10-051-154-724	Sequence 724, App
28	1052	41.1	408	12	US-10-051-999-199163	Sequence 199163, A
29	1022	39.9	417	12	US-10-051-114-49121	Sequence 49121, A
30	780.5	30.5	336	12	US-10-051-114-63429	Sequence 63429, A
31	700	27.4	249	12	US-10-051-999-130445	Sequence 130445, A
32	685	26.8	231	12	US-10-051-114-50887	Sequence 50887, A
33	678	26.5	487	9	US-09-795-693-27	Sequence 27, Appl
34	678	26.5	487	13	US-10-051-139-14	Sequence 14, Appl
35	678	26.5	487	14	US-10-051-239-27	Sequence 27, Appl
36	678	26.5	487	14	US-10-199-485-27	Sequence 27, Appl
37	678	26.5	487	15	US-10-391-399-45	Sequence 45, Appl
38	668	26.1	457	15	US-10-369-493-23324	Sequence 23324, A
39	667.5	26.1	486	9	US-09-860-232A-7	Sequence 7, Appl
40	664.5	26.0	488	12	US-09-794-822-11	Sequence 11, Appl
41	664.5	26.0	488	13	US-10-094-059-4	Sequence 4, Appl
42	664.5	26.0	488	14	US-10-170-528-5	Sequence 5, Appl
43	664.5	26.0	488	14	US-10-162-012-46	Sequence 46, Appl
44	664.5	26.0	488	14	US-10-062-960B-4	Sequence 4, Appl
45	664.5	26.0	488	14	US-10-144-624-4	Sequence 4, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-051-902-20  
; Sequence 20, Application US/10051902  
; Publication No. US20020178468A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingney, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/10/051,902  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 20  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-051-902-20

Query Match	100.0%	Score 2559	DB 13	Length 513
Best Local Similarity	100.0%	Pred. No. 1.7e-230		
Mismatches	513	Conservative	0	Mismatches 0; Indels 0; Gaps 0;
Qy	1	MASDELAKAVEPRKGNVYKASICAILASMASVILGYDGVMSGAAMYIKDLNTDVQL	60	
Db	1	MASDELAKAVEPRKGNVYKASICAILASMASVILGYDGVMSGAAMYIKDLNTDVQL	60	
Qy	61	ELIIGILSYLSFGSFAGARTSDRIGRLTVVFAAVIFVGSLLMGFAVNYGMAGRFV	120	
Db	61	ELIIGILSYLSFGSFAGARTSDRIGRLTVVFAAVIFVGSLLMGFAVNYGMAGRFV	120	
Qy	121	AGVGVGYGMIAPVYTAISPAASGFLTTPEVFINTIGILYLSNFAFARPLHLGWR	180	
Db	121	AGVGVGYGMIAPVYTAISPAASGFLTTPEVFINTIGILYLSNFAFARPLHLGWR	180	

QY 181 VMLAIGAVPSGLLALLVFCMPESPRVLVKGRLADARAVLEKTSATPEEAERLADIKAA 240  
 DB 181 VMLAIGAVPSGLLALLVFCMPESPRVLVKGRLADARAVLEKTSATPEEAERLADIKAA 240  
 QY 241 AGIPKGLDGVVTPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOAGSDSV 300  
 DB 241 AGIPKGLDGVVTPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOAGSDSV 300  
 QY 301 VOYSARLFKSAGITDDNKLIGVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGMIVS 360  
 DB 301 VOYSARLFKSAGITDDNKLIGVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGMIVS 360  
 QY 361 LILCGLSGLTVAGHHPDPTKVAMAVALCIASITLSYIAFFSIGLGPITGVYTSIFPLQVRAL 420  
 DB 361 LILCGLSGLTVAGHHPDPTKVAMAVALCIASITLSYIAFFSIGLGPITGVYTSIFPLQVRAL 420  
 QY 421 GFAVGASNRVTSVISMFTLSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPGRTRTLE 480  
 DB 421 GFAVGASNRVTSVISMFTLSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPGRTRTLE 480  
 QY 481 EMGKLFMPDGTGMAEAEADAADAAKEKVVELPSSK 513  
 DB 481 EMGKLFMPDGTGMAEAEADAADAAKEKVVELPSSK 513

RESULT 2  
 US-10-051-909-20  
 ; Sequence 20, Application US/10051909  
 ; Publication NO. US20020199217A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Steve  
 ; APPLICANT: Helentjaris, Tim  
 ; APPLICANT: Hitz, Bill  
 ; APPLICANT: Kinney, Tony  
 ; APPLICANT: Tingey, Scott  
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins  
 ; FILE REFERENCE: BB1163 US CIP  
 ; CURRENT FILING DATE: 2002-01-17  
 ; PRIOR APPLICATION NUMBER: 60/083,044  
 ; PRIOR FILING DATE: April 24, 1998  
 ; NUMBER OF SEQ ID NOS: 38  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 20  
 ; LENGTH: 513  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-10-051-909-20

Query Match 100.0%; Score 2559; DB 13; Length 513;  
 Best Local Similarity 100.0%; Pred No. 1.7e-230;  
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASDELAKA--AVEPRKKNVYASICAILASMASVILGYDGVNMGAAVYIKKDLNITDVL 60  
 DB 1 MASDELAKA--AVEPRKKNVYASICAILASMASVILGYDGVNMGAAVYIKKDLNITDVL 60  
 QY 61 EILIGILSYLSLFGSFAGARTSDRIGRLTVVFAAVIFVVGSLLMGFVAVNYGMLMAGRFV 120  
 DB 61 EILIGILSYLSLFGSFAGARTSDRIGRLTVVFAAVIFVVGSLLMGFVAVNYGMLMAGRFV 120  
 QY 121 AGVGVGGMIAFVYTAETISPAASRGFLTTPPEVFINIGILGYISNFAFAPLPHLQWR 180  
 DB 121 AGVGVGGMIAFVYTAETISPAASRGFLTTPPEVFINIGILGYISNFAFAPLPHLQWR 180  
 QY 181 VMLAIGAVPSGLLALLVFCMPESPRVLVKGRLADARAVLEKTSATPEEAERLADIKAA 240  
 DB 181 VMLAIGAVPSGLLALLVFCMPESPRVLVKGRLADARAVLEKTSATPEEAERLADIKAA 240  
 QY 241 AGIPKGLDGVVTPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOAGSDSV 300  
 DB 241 AGIPKGLDGVVTPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOAGSDSV 300

QY 301 VOYSARLFKSAGITDDNKLIGVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGMIVS 360  
 DB 301 VOYSARLFKSAGITDDNKLIGVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGMIVS 360  
 QY 361 LILCGLSGLTVAGHHPDPTKVAMAVALCIASITLSYIAFFSIGLGPITGVYTSIFPLQVRAL 420  
 DB 361 LILCGLSGLTVAGHHPDPTKVAMAVALCIASITLSYIAFFSIGLGPITGVYTSIFPLQVRAL 420  
 QY 421 GFAVGASNRVTSVISMFTLSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPGRTRTLE 480  
 DB 421 GFAVGASNRVTSVISMFTLSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPGRTRTLE 480  
 QY 481 EMGKLFMPDGTGMAEAEADAADAAKEKVVELPSSK 513  
 DB 481 EMGKLFMPDGTGMAEAEADAADAAKEKVVELPSSK 513

RESULT 3  
 US-10-051-902-28  
 ; Sequence 28, Application US/10051902  
 ; Publication NO. US20020178468A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Steve  
 ; APPLICANT: Hitz, Bill  
 ; APPLICANT: Kinney, Tony  
 ; APPLICANT: Tingey, Scott  
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins  
 ; FILE REFERENCE: BB-1163  
 ; CURRENT APPLICATION NUMBER: US/10/051,902  
 ; CURRENT FILING DATE: 2002-01-17  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 28  
 ; LENGTH: 529  
 ; TYPE: PRT  
 ; ORGANISM: Triticum aestivum  
 US-10-051-902-28

Query Match 74.5%; Score 1906.5; DB 13; Length 529;  
 Best Local Similarity 73.8%; Pred. No. 2e-169;  
 Matches 373; Conservative 58; Mismatches 71; Indels 5; Gaps 3;  
 QY 1 MASDELAK--AVEPRKKNVYASICAILASMASVILGYDGVNMGAAVYIKKDLNITDV 58  
 DB 19 MASDELAK--AVEPRKKNVYASICAILASMASVILGYDGVNMGAAVYIKKDLNITDV 78  
 QY 59 QLEILIGILSYLSLFGSFAGARTSDRIGRLTVVFAAVIFVVGSLLMGFVAVNYGMLMAGR 118  
 DB 79 QLEILIGILSYLSLFGSFAGARTSDRIGRLTVVFAAVIFVVGSLLMGFVAVNYGMLMAGR 138  
 QY 119 FVAGVGVGGMIAFVYTAETISPAASRGFLTTPPEVFINIGILGYISNFAFAPLPHLQ 178  
 DB 139 FVAGVGVGGMIAFVYTAETISPAASRGFLTTPPEVFINIGILGYISNFAFAPLPHLQ 198  
 QY 179 WRVMAIGAVPSGLLALLVFCMPESPRVLVKGRLADARAVLEKTSATPEEAERLADIK 238  
 DB 199 WRVMAIGAVPSGLLALLVFCMPESPRVLVKGRLADARAVLEKTSATPEEAERLADIK 258  
 QY 239 AAAGIPKGLDGVVTPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOAGSD 298  
 DB 259 AAAGIPKGLDGVVTPGKEGGGELQVWKLLILSPTPAVRRIILLSAVGLHFFQOAGSD 317  
 QY 299 SVVOYSARLFKSAGITDDNKLIGVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGM 358  
 DB 318 SVVOYSARLFKSAGITDDNKLIGVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGM 377  
 QY 359 VSLICLGSGLTVAGHHPDPTKVAMAVALCIASITLSYIAFFSIGLGPITGVYTSIFPLQVR 418  
 DB 378 VSLICLGSGLTVAGHHPDPTKVAMAVALCIASITLSYIAFFSIGLGPITGVYTSIFPLQVR 437



; APPLICANT: Allen, Steve  
; APPLICANT: Helencjaris, Tim  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB1163 US CIP  
; CURRENT APPLICATION NUMBER: US/10/051,909  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 60/083,044  
; PRIOR FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 22  
; LENGTH: 510  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; NAME/KEY: UNSURE  
; LOCATION: (102)  
US-10-051-909-22

Query Match 73.2%; Score 1872.5; DB 13; Length 510;  
Best Local Similarity 72.7%; Pred. No. 2.9e-166; Mismatches 60; Indels 3; Gaps 2;  
Matches 372; Conservative 60; Mismatches 77; Indels 3; Gaps 2;  
QY 1 MASDELAKEVEPRKKNVYASICALASMASVILGYDGVMSGAAMYIKKDLNITDVL 60  
DB 1 MASAAPEAVAPKKNVRFACALLASMTSILLGYDGVMSGASLYIKKDFNISDGV 60  
QY 61 EILIGLSYLSFGSPAGARTDGRRLTVFAVAVFVGSLLGFAVNYCMLMAGRFV 120  
DB 61 EVLMGLNLSYLSFGSPAAGRTSDWIGRRYTVFAVAVFVGSLLGFAVNYCMLMAGRFV 120  
QY 121 AGVGVGYGMIAPVYTAETSPAASRGFLTFPEVFINIGLLGYLSNFAFARLPHLGWR 180  
DB 121 AGIGVYALMIAPVYTAETSPAASRGFLTFPEVFINIGLLGYLSNFAFARLPHLGWR 180  
QY 181 VMLAIGAVPSGLLALVFCMPSPRWLVKGLADARAVLEKTSATPEAERLADIKAA 240  
DB 181 IMLGIAAPSLLALVLCMPSPRWLVKGLADARAVLEKTSATPEAERLADIKAA 240  
QY 241 AGIPKGLDGVTVVPGKEGGELQVWKLLISPTPAVERILLSAVGLHFFQAGSDSV 300  
DB 241 AGIPEELDGVTVVPGKEGGELQVWKLLISPTPAVERILLSAVGLHFFQAGSDSV 300  
QY 301 VOYSAELFXSAGITDDNKLGLVTCVAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360  
DB 301 VVYSPVLVFXSPGLTNDKHLGTTWPGVTKRLFIILLATFIDGVGRRPLLLISTGGMIVS 360  
QY 361 LICLSGLTVAGHPPTKVAWALCIASLTSLYIAFFSICLGPITGVTSIFFPQLQVRL 420  
DB 361 LICLSGLTVAGHPPTKVAWALCIASLTSLYIAFFSICLGPITGVTSIFFPQLQVRL 420  
QY 421 GFVAVGASNRVTSVAVISMTFSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPSTRGRLE 480  
DB 420 GCSLGVANRVTSVAVISMTFSLSKAITIGGSFFLYSGIAALAWVFFYTLPTGRLE 479  
QY 481 EMKLFMPDPTGMAEAEADAAKEKVELPSS 512  
DB 480 EMKLFMPDPTGMAEAEADAAKEKVELPSS 509

RESULT 7  
US-10-051-902-26  
; Sequence 26, Application US/10051902  
; Publication No. US20020178468A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/10/051,902  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 26  
; LENGTH: 539  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-10-051-902-26

Query Match 69.0%; Score 1764.5; DB 13; Length 539;  
Best Local Similarity 68.8%; Pred. No. 4.1e-156; Mismatches 63; Indels 5; Gaps 2;  
Matches 351; Conservative 63; Mismatches 91; Indels 5; Gaps 2;  
QY 6 LAKAVEPRKKNVYASICALASMASVILGYDGVMSGAAMYIKKDLNITDVLQLEILIG 65  
DB 31 LPAVPEPKKGNVRFACALLASMTSILLGYDGVMSGASLYIQKDLKINDTQLEVLWG 90  
QY 66 ILSLSLFGSPAGARTSDRIGRLTVFAVAVFVGSLLGFAVNYCMLMAGRFVAGVGV 125  
DB 91 ILNVYSLIGSFAAGRTSDWIGRRYTVFAVAVFVGSLLGFAVNYCMLMAGRFVAGVGV 150  
QY 126 GYGMIAPIVYTAETSPAASRGFLTFPEVFINIGLLGYLSNFAFARLPHLGWRVWLAI 185  
DB 151 GYALMIAPIVYTAETSPAASRGFLTFPEVFINIGLLGYLSNFAFARLPHLGWRVWLAI 210  
QY 186 GAVPSGLLALLVFCMPSPRWLVKGLADARAVLEKTSATPEEAAERLADIAAGIPL 245  
DB 211 GAVPSGLLALLVFCMPSPRWLVKGLADARAVLEKTSATPEEAAERLADIAAGIPL 270  
QY 246 GLDGVTVVPGKEGGELQVWKLLISPTPAVERILLSAVGLHFFQAGSDSVVOYSA 305  
DB 271 GLDGVTVVPGKEGGELQVWKLLISPTPAVERILLSAVGLHFFQAGSDSVVOYSA 330  
QY 306 RLPKSAGITDDNKLGLVTCVAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSICLG 365  
DB 331 LVPSKAGITDGSRLGTTVAVGATNTVFILVATFLLDRIRRRPLVLTSTGGMLSVGLA 390  
QY 366 SGLTVAGHPPTKVAWALCIASLTSLYIAFFSICLGPITGVTSIFFPQLQVRL 425  
DB 391 TGLTVISRHPDEKITWAILVLCIFCINAYVAFSSIGLGPITWVYSSIFFPLHVRALGCSLG 450  
QY 426 VASNRVTSVAVISMTFSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPSTRGRLEBMGKL 485  
DB 451 VAVNRVTSVAVISMTFSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPSTRGRLEBMGKL 510  
QY 486 FGMPDPT---GMAEAEADAAKEKVELPSS 512  
DB 511 FGNTATHKQGAEEADDDAG--EKKVEMAAT 538

RESULT 8  
US-10-051-909-26  
; Sequence 26, Application US/10051909  
; Publication No. US20020199217A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Helencjaris, Tim  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB1163 US CIP  
; CURRENT APPLICATION NUMBER: US/10/051,909  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 60/083,044  
; PRIOR FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 26

```

; LENGTH: 539
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-909-26

Query Match          69.0%; Score 1764.5; DB 13; Length 539;
Best Local Similarity 68.8%; Pred. No. 4.1e-156;
Matches 351; Conservative 63; Mismatches 91; Indels 5; Gaps 2;

QY 6 LAKAVEPRKGNVYKASICAILASMASVILGYDVGMSGAAMYIKKDLNITDVLQLEILIG 65
DB 1 LPAAVEPKKGNVRFAPACAILASMTSILLGYDVGMSGASLYIQKDKINDTQLEVLWG 90
QY 66 ILSLSLFGSFAGARTSDRIGRLTVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGV 125
DB 91 ILNVYSLGSPAGRTSDWIGRRFIVFAAVIFFFAGALIMGFSVNYAMLMGFRFVAGIGV 150
QY 126 GYGMIAPIVYTAETISPAASRGELTTFPEVFINIGILGYLSNFAFARLPLHLGWRVMLAI 185
DB 151 GYALMIAPVNTGEVSPASARGVLTSPPEVFINIGILGYLSNFAFARLPLHLGWRVMLGI 210
QY 186 GAVPSGLLALVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAAAGIPK 245
DB 211 GAVPSVLLAFVLGMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAAAGIPL 270
QY 246 GLDGDWTVPGKEGGGELQVWKLLISPTPAVRRIILLSAVGLHFFFOAGSGSDSVVQYSA 305
DB 271 GLDGDWVPVKNGSGSEKRVLKDLISPTTAMRHILAGIGHFFFOAGSGSDIDAVLYSP 330
QY 306 RLFSAGITDDNKLGLVTCVGVTKTFFILVATFLLDRAGRRPMLLISLSTGMIYSLICLGS 365
DB 331 LVFSAGITDGRSLRGTTTAVGAINTVFVLVATFLLDRAGRRPMLLISLSTGMIYSLICLGS 390
QY 366 SGLTVAGHPPTKVAVAVLALCIASLTSYIAFFSIGLGPITGVYTSFIFPLOVRALGFAVG 425
DB 391 TGLTVISHPEDEKITWAILVLCIFCMAYAVFSIGLGPITGVYTSFIFPLOVRALGCSLG 450
QY 426 VASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRITLLEMGKL 485
DB 451 VAVNRLTSGVISMFTISLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRITLLEMGKL 510
QY 486 FGMPDT---GMAEEAEADAARKEKVELPSS 512
DB 511 FGNATHKQGAEEADDDAG--EKKVEMAAT 538

RESULT 9
US-10-425-114-68399
; Sequence 68399, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68399
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73004E08_FLI.pep
US-10-425-114-68399

Query Match          63.5%; Score 1626; DB 12; Length 327;
Best Local Similarity 93.4%; Pred. No. 1.8e-143;
Matches 326; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 1 MASDELAKAVEPRKGNVYKASICAILASMASVILGYDVGMSGAAMYIKKDLNITDVLQ 60
DB 11 MASDELAKIVEPRKGNVYKASICAILASMASVILGYDVGMSGAAMYIKKDLNITDVLQ 70
QY 61 EILIGILSLYSLFGSFAGARTSDRIGRLTVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
DB 71 EILIGILSLYSLFGSFAGARTSDRIGRLTVFAAVIFFFAGSLLMGFAVNYGMLMAGRFV 130
QY 121 AGVGVGYGGMIAPIVYTAETISPAASRGELTTFPEVFINIGILGYLSNFAFARLPLHLGWR 180
DB 131 AGVGVGYGGMIAPIVYTAETISPAASRGELTTFPEVFINIGILGYLSNFAFARLPLHLGWR 190
QY 181 VMLTAGVPSGLLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAA 240
DB 191 VMLTAGVPSALLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAA 250
QY 241 AGIPKGLDGDVVTVPGK-EQGGELQVWKLLISPTPAVRRIILLSAVGLHFFFOAGSGSDS 299

Matches 325; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 187 AVPSGLLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAAAGIPKG 246
DB 1 AVPSGLLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAAAGIPKG 60
QY 247 LDGDWTVPGKEGGGELQVWKLLISPTPAVRRIILLSAVGLHFFFOAGSGSDSVVQYSA 306
DB 61 LDGDWTVVPNKKEGGGELQVWKLLISPTPAVRRIILLSAVGLHFFFOAGSGSDSVVQYSA 120
QY 307 LFSAGITDDNKLGLVTCVGVTKTFFILVATFLLDRAGRRPMLLISLSTGMIYSLICLGS 366
DB 121 LFSAGITDDNKLGLVTCVGVTKTFFILVATFLLDRAGRRPMLLISLSTGMIYSLICLGS 180
QY 367 GLTVAGHPPTKVAVAVLALCIASLTSYIAFFSIGLGPITGVYTSFIFPLOVRALGFAVG 426
DB 181 GLTVAGHPPTKVAVAVLALCIASLTSYIAFFSIGLGPITGVYTSFIFPLOVRALGFAVG 240
QY 427 ASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRITLLEMGKL 486
DB 241 ASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRITLLEMGKL 300

RESULT 10
US-10-425-114-61926
; Sequence 61926, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61926
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3593-006-P8_FLI.pep
US-10-425-114-61926

Query Match          63.0%; Score 1612.5; DB 12; Length 356;
Best Local Similarity 94.2%; Pred. No. 3.8e-142;
Matches 326; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 1 MASDELAKAVEPRKGNVYKASICAILASMASVILGYDVGMSGAAMYIKKDLNITDVLQ 60
DB 11 MASDELAKIVEPRKGNVYKASICAILASMASVILGYDVGMSGAAMYIKKDLNITDVLQ 70
QY 61 EILIGILSLYSLFGSFAGARTSDRIGRLTVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
DB 71 EILIGILSLYSLFGSFAGARTSDRIGRLTVFAAVIFFFAGSLLMGFAVNYGMLMAGRFV 130
QY 121 AGVGVGYGGMIAPIVYTAETISPAASRGELTTFPEVFINIGILGYLSNFAFARLPLHLGWR 180
DB 131 AGVGVGYGGMIAPIVYTAETISPAASRGELTTFPEVFINIGILGYLSNFAFARLPLHLGWR 190
QY 181 VMLTAGVPSGLLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAA 240
DB 191 VMLTAGVPSALLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIKAA 250
QY 241 AGIPKGLDGDVVTVPGK-EQGGELQVWKLLISPTPAVRRIILLSAVGLHFFFOAGSGSDS 299
```

Db 251 AGIPDLGDGVAVVPSKGRDSSSETQVWKQILSPTFAVRILLSSAVGLHFFQOASGSDS 310  
QY 300 VVOYSARLFKSGAGITDDNKLGLGVTCAGVTKTFFILVATPLLDRA 345  
Db 311 VVOYSARLFKSGAGITDDNKLGLGVTCAGVTKTFFILVATPLLDRA 356  
RESULT 11  
US-10-425-114-39509  
; Sequence 39509, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313) B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; PRIOR FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 39509  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700239390\_Fli.pgp  
US-10-425-114-39509  
Query Match 62.4%; Score 1597; DB 12; Length 380;  
Best Local Similarity 85.4%; Pred. No. 1.2e-140;  
Matches 322; Conservative 21; Mismatches 28; Indels 6; Gaps 4;  
QY 139 ISPAASRGFLTPPEVFINIGILLGYSNFAFARLPLHLGWRVMAIGAIVPSGLLALV 198  
Db 1 ISPAASRGFLTPPEVFINIGILLGYSNFAFARLPLRFGWRVMLGIGAAPAGLILALMV 60  
QY 199 CMPEPRHLVKGRLADARAVLEKTSATPEBAERLADIKAAAGPKGLGDGVTVPGKE 258  
Db 61 WPESPRHLVKGRLADARAVLEKTSASPEEAERLADIKAAAGPKGLGDGVTVAVP-KE 119  
QY 259 QGSGELQVWKKLILSPTFAVRILLSSAVGLHFFQOASGSDSVVOYSARLFKSAGITDDNK 318  
Db 120 RSGGEMQVWRELILSPTFAVRILLSSAVGLHFFQOASGSDSVVLYSPVFKSAGITDDNK 179  
QY 319 L--LGVTCAVGVTKTFILVATPLLDRAGRRLILLISTGGMIVSLICLGSGLTVAGHHPD 376  
Db 180 LLDLGVTCAGVAKTLFIPVATPLLDRAGRRLILLISTGGMIVSLVGLGTGLTVVGHHPD 239  
QY 377 TKVAMAVALCIATLSVIAFFSTGLGPITGVYTSFIFPQVRALGFVAVGASNRVTSVI 436  
Db 240 AKTPSAVALCIATLSVIAFFSTGLGPITGVYTSFIFPQVRALGFVAVGASNRVTSVI 299  
QY 437 SMTFLSKAITIGGSFFLYSIAAVAVVFFFTCLPSTRGRLEEMKLFQM--PDTGMA 494  
Db 300 SMTFLSKAITIGGSFFLYSIAAVAVVFFFTCLPSTRGRLEEMKLFQMPEPDTDMA 359  
QY 495 EEAEADAAKKEVLEPS 511  
Db 360 -EAYNAAKKEVEMPA 375  
RESULT 12  
US-10-051-902-24  
; Sequence 24, Application US/10051902  
; Publication No. US20000178488A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/10/051,902  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 24  
; LENGTH: 523  
; TYPE: PRT  
; ORGANISM: Glycine max  
; US-10-051-902-24  
Query Match 60.3%; Score 1542; DB 13; Length 523;  
Best Local Similarity 62.0%; Pred. No. 2.6e-135;  
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;  
QY 1 MASDELAKAVE-----PRKGNVYKASICAILASNASVILCYDGVMSGAMVYK 50  
Db 1 MTEGKLVAAEAHKTLDPPPKKRNKYAFACAMLASMTSILLGYDGVMSGAAIYK 60  
QY 51 KDLNITDVOLEILIGILSYSLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVN 110  
Db 61 RDLKVSDEQLEILLGIINLYSLGSLAGRTSDWIGRTYIVFAGTFFVGGALLMGFSN 120  
QY 111 YGMLMAGRFVAGVGVGGMIAPVYTAETSPAASRGFLTPPEVFINIGILLGYSNPAF 170  
Db 121 YSFLMFGRFVAGIGIGYALMIAPVYTAESPASSRGFLTSFPEVINGGILIGYSNPAF 180  
QY 171 ARPLHLGWRVMAIGAIVPSGLLALVFCMPESPRHLVKGRLADARAVLEKTSATPEBA 230  
Db 181 SKLTLYGVEMVILGVAIPSVLLTVGLVAMPESPRHLVKGRLADARAVLEKTSATPEBA 240  
QY 231 AERLADIKAAAGPKGLGDGVTVPGKEQGGBLQVWKKLILSPTFAVRILLSSAVGLH 290  
Db 241 QLRALAEIKQAAGPESCNDDVVQVKNQSGEG---VWKEFLYPTAIRHIVIAALGIHF 297  
QY 291 FQOASGSDSVVOYSARLFKSAGITDDNKLIGVTCAVGVTKTFILVATPLLDRAGRPL 350  
Db 298 FQOASGSDSVVOYSARLFKSAGITDDNKLIGVTCAVGVTKTFILVATPLLDRAGRPL 357  
QY 351 LITSGGMIVSLICLGSGLTVAGHHPDTKVAVALCIATLSVIAFFSTGLGPITGVYTS 410  
Db 358 LSSVGGMWLSLLTALSLTVI-DHSERKLMWAVGSSIAMVLAIVATFSGAGITWVYS 416  
QY 411 EIPFLQVRALGFVAVGASNRVTSVIAFFSTGLGPITGVYTSVIAFFSTGLGPITGVYTS 470  
Db 417 EIPFLQVRALGFVAVGASNRVTSVIAFFSTGLGPITGVYTSVIAFFSTGLGPITGVYTS 476  
QY 471 LPETRGRTLEEMKLF 487  
Db 477 LPETRGRTLEEMKLF 493  
RESULT 13  
US-10-051-909-24  
; Sequence 24, Application US/10051909  
; Publication No. US20020199217A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Helentjaris, Tim  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB1163 US CIP  
; CURRENT APPLICATION NUMBER: US/10/051,909  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 60/083,044  
; PRIOR FILING DATE: April 24, 1998



```

; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-909-24

Query Match      60.3%; Score 1542; DB 13; Length 523;
Best Local Similarity 62.0%; Pred. No. 2.6e-135;
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKGNVYASICAILASMASVILGYDVGWMSAAAYIK 50
DB 1 MTEGKLVEAAEAKHTLQDPDPKKRKNKYAPACAWLASMTSILGYDVGWMSAAAYIK 60

QY 51 KDLNITDVLEILGILSYLFGSFAGARTSDRIGRELTVVFAAVIFFVGSLLMGPAVN 110
DB 61 RDLXGDEQIEILLGIINLYSLIGSCLAGRTSDWIGPRYTVFAGTFFVYGALLMGFSPN 120

QY 111 YGMLMAGRFVAGVGYGGMIAPVYTAELSPAASRGFLTTPEVFINIGILLGYSNFAF 170
DB 121 YSFLMFRFVAGIGIGYALMIAPVYTAEVSPAASSRGFLTSPEVFINIGILGYSNFAF 180

QY 171 ARFLPHLGMRVLMIAIGAVPSGLLAILVFCMPSPRWLVKGLADARAVLEKTSATPEEA 230
DB 181 SKLTLVGWRMMLGVGAISVLLTVGVLAIMPESPRWLVMRGSLGEARKVLNKTSDSKEEA 240

QY 231 AELADIKAAGIPKGLGDVTVTPGKQGGELQVWKKLILSTPAVRRIILSAVGLHF 290
DB 241 QLSLAIRKQAAGIPESCNDDVVQVKNQSGEG---VMKELFLYPTPAIRHIVIAALGIHF 297

QY 291 FQAQSGDSVVQYSARLFKSAGITDDNKLGLVTCAGVYTKTFFILVATFLLDRAGRPLL 350
DB 298 FQAQSGVDVLYSPRIPEKAGITNDTHKLATVAVGFKTVFILAATFTLDRVGRPRL 357

QY 351 LISTGNIVSLICLGSGLTVAGHPDPTKVAVAVALCIASLTLSYIAFFSIGLGPITGYTS 410
DB 358 LSSVGMWLSLTLTAISLTVI--DHSEKLMWAGSSIAMVLAYVATFSICAGPITWYSS 416

QY 411 EIPFLQVRALGFVAGVNASNRVTSAVISMTFLSKAITGSPFLYSGIAAVAVVFFTC 470
DB 417 EIPFLKRAQGAAGAAVAVNRITTSVVVSMTFLSLTRAITIGGAFFLYCGIATVGWIFFTV 476

QY 471 LPETRGTLLEMGKLF 487
DB 477 LPETRGKTLDEMGSEFG 493

RESULT 14
US-10-425-114-56035
; Sequence 56035, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56035
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701204606_FLI pep
US-10-425-114-56035

```

```

Query Match          58.1%; Score 1487.5; DB 12; Length 517;
Best Local Similarity 62.3%; Pred.No. 3.2e-130;
Matches 297; Conservative 69; Mismatches 104; Indels 7; Gaps 3;

QY      11 EPRKKGNKVASICAILASHMASVTLGYDVGWMSGAAMYIKKDNINITDVQLEILIGILSLY 70
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      17 KPRN---KAFACAILASWTSLILGYDVGWGAALYIQDLKVSVDVQLEILINGINLY 73
QY      71 SLFGSFAGATSDRIGRRLTVFAAVIFFVGSLLMGFAVNVGMAGRFVAGVGVGGM 130
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      74 SPVGSFIAGRTSDWIGRRTIVLAGAIFFFVGAILMGFSPNYAFIMFGQFPFAGVIGGFAFL 133
QY      131 IAPVYTAIEISPAASRGFLTTFPEVFNINIGLLGVLSNFAFARLPLHLGWRVMLAIGAVPS 190
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      134 IAPVYSEISPSRSRGFLTSLPEVFLNGGLIGYISNYGFSKLALRLGNWRLMLGVGAIPS 193
QY      191 GLIALLVFCMPESPRMLVLKGRADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGD 250
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      194 ILIGVAVLAMPESPRLVAKRGLGEAKRVLVYKISEEAEARLEADIKDTAGTPQCDDDD 253
QY      251 VVTVPQKEQGGELQWVKLILSPTPAVRRIILASVGLHFFQOASGSDSVVQYSARLFKS 310
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      254 VLVSKQTHG---VWRELFLHPTPAVRHIFASIGIHFFAQTGIDAVVLVSPRIFEX 310
QY      311 AGITDNKLLGVTCVAVGVTKTFIFLVATFLDRAGRPLLLISTGGMVSLICIGSLTV 370
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      311 AGIKSDNYRLAVTAVGVFKTVSILVATPFLDRAGRVRLLLCSVGLISLLTGLSLTV 370
QY      371 AGHPHDTKVAAWVALCIIASTLSYAFSTIGLGPITGVTVSEIFPQVLRALGFVAGVASNR 430
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      371 V-DHSQTLNWAAGLSIAVLSVATFSTIGSGPITVWYSEIFPFLRLAQGVGAIVANR 429
QY      431 VTSAVISMFTLSUSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRTLEEMKFLG 487
      Db      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      430 VTSGVIAFTLSLQKAITIGGAPFLFAGVAAVAWIEHYTLLPETRGKTLEETKESFG 486

RESULT 15
US-09-774-381-40
; Sequence 40, Application US/09774381
; Publication No. US20030082677A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Pan, Yang
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NOVEL EDJRF, MTR-1, LSP-1, TAP-1, AND PA-I MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MN1-107CP2
; CURRENT APPLICATION NUMBER: US/09/774,381
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 08/941,354
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/010,674
; PRIOR FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: 60/061,149
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/014,347
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 60/061,159
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/474,151
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 09/004,206
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/061,143
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/483,414
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 09/213,571
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,890

```

```
/ PRIOR FILING DATE: 1997-12-19
/ NUMBER OF SEQ ID NOS: 59
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 40
/ LENGTH: 548
/ TYPE: PRT
/ ORGANISM: Beta vulgaris
US-09-774-381-40

Query Match      58.0%; Score 1483.5; DB 10; Length 548;
Best Local Similarity 61.6%; Pred. No. 8.3e-130;
Matches 299; Conservative 63; Mismatches 112; Indels 11; Gaps 5;

QY 4 DELAKAVEPRKGNVYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVOLEIL 63
DB 24 DELKK---PPKRN--KFAPACATLASMTSVLLGYDIGVMSGAIYLLKEDWHISTQIGVL 78
QY 64 IGLSLYSLFGSFAGARTSDRIGRELTVVFAAVIFFVGSLLMGFAVNYGMIMAGRFRVAGV 123
DB 79 VGILNIYCLFGSFAAGRTSDWIGRRYTIIVLAGAIFVVGALLMGFATNYAFLMVGRFVTGI 138
QY 124 GVGYGMIAPVYTAISPAASRGFLTPPEVFINIGILLGYLSNFAPARLPHILGWRVWL 183
DB 139 GVGVALMIAPVYTAEVSPASSRGFLTPPEVFINAGILLGYLSNLAFLSSPLTHLSWRFML 198
QY 184 AIGAVPSGLLALLVFCMPSPRWLVKGRLDARAVLEKTSATPEEAERLADIKAAAGI 243
DB 199 GIGALPSIFLAGVLPAMPSPRWLVMOGFLGDAKKVLRISDSPEAQRLSEIKQTAGI 258
QY 244 PKGLDGVVTVPGKEGGGELQVWKXLIILSPTPAVRILLSAVGLHFTQQASGSDSVVQY 303
DB 259 PAECDEDIYKVEKTIKSGN-AWVKELFFENPTFAVRRAVIAGIGIHFTQQASGIDAVVLY 317
QY 304 SARLFKSAGITDDNKLIGVTCAGVTKTPEILVATFLDDRGRRLILLISTGGMIVSLIC 363
DB 318 SPRIFOSAGITNARKOLLATVAVGVVKTFLFVATFQLDKYGERPLLLTSVSGMIILIT 377
QY 364 LGSGLTVA--GHRPDTKVAVAVALCIASLTSLYAFPSIGLGPITGVYTSEIFPQVRLG 421
DB 378 LAMSLTVIDHSHH---KITWATALCITMVCVAVASFSIGLGPITWVYSSEVPFLRLAQG 434
QY 422 FAVGVASNRVTSNVTSLISLKAITIGGSFELYSGIAAVAVVFFETCLPETRGRTLEE 481
DB 435 TSMGVAVNRVVGVSIFIFLPLSHKITTGGAFFLFGGIAIIAFFFLLFPLPETRGRTLEN 494
QY 482 MGKLF 486
DB 495 MHLP 499
```

Search completed: June 30, 2004, 18:37:03  
Job time : 51 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:30:53 ; Search time 21 Seconds  
(without alignments)  
2349.822 Million cell updates/sec

Title: US-10-051-902A-20

Perfect score: 2559

Sequence: 1 MASDELAKAVEPRKKNVKY.....AEEEDAAAKKVVLPSSK 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1483.5	58.0	549	2	T14606
2	1479.5	57.8	511	2	H84536
3	1463.5	57.2	511	2	A84537
4	1346	52.6	493	2	A85433
5	1316	51.4	508	2	G84564
6	1176	46.0	547	2	C84593
7	668	26.1	457	2	E70070
8	629.5	24.6	580	2	D86426
9	624.5	24.4	461	2	D70073
10	606	23.7	521	2	G84864
11	597	23.3	582	2	F71431
12	595	23.3	560	2	T31485
13	588	23.0	517	2	D86539
14	586.5	22.9	472	2	B26430
15	586.5	22.9	472	2	B31091
16	586.5	22.9	472	2	B85936
17	586	22.9	473	2	G69789
18	586	22.9	557	2	T38125
19	579	22.6	472	2	S47089
20	575.5	22.5	471	2	A80868
21	574.5	22.5	487	2	A80868
22	572.5	22.4	580	2	D84772
23	568.5	22.2	482	2	B69803
24	568	22.2	502	2	B70845
25	565	22.1	584	2	S69555
26	564	22.0	606	2	T27072
27	563	22.0	612	2	B40538
28	553	21.6	464	2	C91106
29	553	21.6	464	2	F85951

30 552.5 21.6 639 2 T23658  
31 547 21.4 464 2 F65079  
32 546 21.3 464 2 AC0877  
33 546 21.3 575 2 T43400  
34 544 21.3 491 2 A26430  
35 544 21.3 491 2 F91255  
36 544 21.3 491 2 F86096  
37 538 21.0 464 2 F69587  
38 526.5 20.6 613 2 T27077  
39 526 20.6 468 2 S10014  
40 523 20.4 507 2 T01844  
41 522 20.4 433 2 G86812  
42 522 20.4 534 2 S38435  
43 520 20.3 516 2 T12199  
44 520 20.3 547 2 A48442  
45 515.5 20.1 522 2 E86246

#### ALIGNMENTS

##### RESULT 1

T14606

probable sugar transport protein 205 - beet

C;Species: Beta vulgaris (beet)

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 01-Dec-2000

C;Accession: T14606; T14617

R;Chou, T.J.; Bush, D.R.

Plant Physiol. 110, 511-520, 1996

A;Title: Molecular cloning, immunochemical localization to the vacuole, and expression i

A;Reference number: Z18131; MUID:96351183; PMID:8742332

A;Accession: T14606

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-549 <CHI>

A;Cross-references: EMBL:U64902; NID:gl778092; PID:gl778093

A;Accession: T14617

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-541, 'SVQV' <CH2>

A;Cross-references: EMBL:U64903; NID:gl778094; PID:gl778095

C;Genetics:

A;Note: BvcdNA-205; BvcdNA-397

C;Superfamily: Glucose transport protein

C;Keywords: sugar transport

Query Match 58.0%; Score 1483.5; DB 2; Length 549;

Best Local Similarity 61.8%; Pred.No. 9.8e-101;

Matches 299; Conservative 63; Mismatches 112; Indels 11; Gaps 5;

QY 4 DELAKAVEPRKKNVKYASICAILASMASVILGYDGVMSGAAMYIKKDLNITDVLLEIL 63  
DB 25 DPLKK--PPKRN--KFAFACATLASMTSVLLGYDGVMSGAIYUKEDWHISDTQIGVL 79  
QY 64 IGILSYSLFGSPAGARTSDRGRLLTVVFAAVIFVGSLLMGFAVNYGMLMAGRIVAGV 123  
DB 80 VGLINTYCFGSGAAGTSDWIGRRYTVIAGAFVFGALLMGFATNYAFLVMGFRVTGI 139  
QY 124 GVGYGMIAPVVTAEISPAASRGFTTFPEVINTIGILLYLSNFAFARLPHLHGRVML 183  
DB 140 GVGYALMIAPVVTAEVSPASSRGFTTFPEVINAGILLGYLSNLAFFSLPHLSWRFML 199  
QY 184 AGAVPSGALLALVFCMPSPRWLVKRLADARAVLEKTSATPEEAARLADIKAAGI 243  
DB 200 GIGAFISIFLAGVLAPEPSPRWLVNQGRGLGAKVLRNIDSPEAQRLSEIKQTAGI 259  
QY 244 PKGLDGVVTVPKQGGGLOVWKKLILSPTPVRRILLSAVLGHFFQOAGSSDSVWCY 303  
DB 260 PAECDEDIYKVEKTKISGN-AWKELFPNTPVRRVAVIAGIGHFFQOAGSIDAVVLY 318  
QY 304 SARLPKSGITDDNKLGLVTCVAGTKTFILVATFLDRAGRRLPILLISTGCMIVSLIC 363  
DB 319 SPRIQFSGAGITNARKQLLATVAVGVVKILFVATFQLDKYGRRPRLTLTSVGGMIILIT 378



R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: A85433

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <STO>

A:Cross-references: GB:NC\_001268; NID:97270615; PIDN:CAB80333.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4936670

A:Map position: 4

C:Superfamily: glucose transport protein

Query Match 52.6%; Score 1346; DB 2; Length 493;  
Best Local Similarity 55.5%; Pred. No. 9.6e-91;  
Matches 268; Conservative 94; Mismatches 115; Indels 6; Gaps 5;

QY 6 LAKAVEPKGNV-KYASICAILASMASVILGYDGVMSGAAMYKKDLNITDVQLEILI 64  
DB 1 MADQISGEKPGACVNFALQCAIVASIVSIIFGYDTGVMSGAMVFEEBKTNDVQIEVLT 60  
QY 65 GILSYLSFGSPAGARTSDRIGRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRFVAGV 124  
DB 61 GILNLCALVGLLAGRTSDIIGRRVTIVLASILFVLGSGILMGWGNYPVLLSGRCTAGLG 120  
QY 125 VCYGGMIAPVYTABISPAASRGFLTFPEVINIGILGYLSNFAFAPLPHLGWRVMLA 184  
DB 121 VGFAVMVAPVYSAETATASHRGLASLPHLCISIGILGYLVNFFSKPLPHIGWRMLG 180  
QY 185 IGAVPSGLLALVFCWPSRWLVKGLADARAVLEKTSATPEEAERLADIKAAAGI - 243  
DB 181 IAAVPSVLVAGILKWPSPRWLVKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 240  
QY 244 PKGLDGVTVVPGKQGGGEIQVKKLILSTPTPAVRILLVAVGHHFFQAGSDSVVQY 303  
DB 241 PKCVD-DVVRMEGKTHGE--GVKKELILRPTPAVRLLTALGTHFFQHASGTEAVLLY 297  
QY 304 SARLFKSGITDDNKLGLVTCAGVTKTFILVATFLDRAGRPLILISTGGMIVSLIC 363  
DB 298 GPRIFKAGITTKDLFLVTVGVGMKTTFTALLDLKVRGRKLLTSTVGGWVIALTM 357  
QY 364 LGSGLTVAGHHPDTKVAWALCIASLTLSYIAFFSIGLGPITGVYTSIIFPQVLRALGFA 423  
DB 358 LGFGLTMA-QNAGGLKAWLVLSIAVAFVFAFFSIGLGPITWYSSEVFFLKRAQAS 416  
QY 424 VGVASNRVTSAVISMTFSLSKAITIGSGFFLYSGIAAVAVVFFTCLPETRGTLBGM 483  
DB 417 LGVAVNRVYVATVSNFSLTSAITTGAFPFAGVAAVAVVFFLLPETKGSLEEIE 476  
QY 484 KLF 486  
DB 477 ALF 479

RESULT 5

G84564

probable sugar transporter [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: G84564

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84564

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-508 <STO>

A:Cross-references: GB:AE002093; NID:94218010; PIDN:AAD12218.1; GSPDB:GN00139

C:Genetics:

A:Gene: AT2g18480

A:Map position: 2

C:Superfamily: glucose transport protein

Query Match 51.4%; Score 1316; DB 2; Length 508;

Best Local Similarity 55.6%; Pred. No. 1.5e-88;

Matches 264; Conservative 83; Mismatches 112; Indels 16; Gaps 4;

QY 19 KYASICAILASMASVILGYDGVMSGAAMYKKDLNITDVQLEILIGILSYLSFGSPAG 78  
DB 20 KFAFGCAIVASIIISIFGYDTGVMSGAQIPRDDLKINDTQIEVLGILNLCALVSLTA 79  
QY 79 ARTSDRIGRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRFVAGVGVYGGMIAPVYTA 138  
DB 80 GKTSDVIGREYTTIALSAVIFLVGSLMGYPNYPVLMVGRCIAGVGVGFALMIAPVSA 139  
QY 139 ISPAASRGFLTFPEVINIGILGYLSNFAFAPLPHLGWRVMLAIGAVPSGLLALLVF 198  
DB 140 ISSASHRGFLTSPCLICISIGILGYVSNFCFKLILKLGWRMLGLAAPPSSILAFGIT 199  
QY 199 CMPESPRWLVKGLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGVTVVPGKE 258  
DB 200 RMPESPRWLVQGRLEAKIMVLSNTEEAERERDILTAAEV-----DVTEI--XE 251  
QY 259 QGGGELQ-----VMKKLILSTPTPAVRILLVAVGHHFFQAGSDSVVQYSARLEKSA 312  
DB 252 VGGVKKXKNGKSVKRELVIKPRPAVRLLIAAVGHHFFEHATGIEAVVLYSPRIFKAG 311  
QY 313 ITDENKILGVTCAGVTKTFILVATFLDRAGRPLILISTGGMIVSLICLSGLTVAG 372  
DB 312 VSKDKLLATVGLTKAFFIILATFLDKVGRKLLTSTGVMFALTSLAVSLTWQ 371  
QY 373 HHPTKVANAVALCIASLTLSYIAFFSIGLGPITGVYTSIIFPQVLRALGFAVAVNRVT 432  
DB 372 RF--GGLAWLSLSIVSTVAFVAFVFFSIGLGPITWYSSEIIFPLRLRAQAGASIGAVNRIM 429  
QY 433 SAVISMTFSLSKAITIGSGFFLYSGIAAVAVVFFTCLPETRGTRLEEMKLF 487  
DB 430 NATVMSFSLSKAITIGGVFFVFAVAGIAVAAWVFFFLPETKGLPEEMKLF 484

RESULT 6

C84593

probable sugar transporter [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 03-Jun-2002

C:Accession: C84593

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84593

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-547 <STO>

A:Cross-references: GB:AE002093; NID:94454470; PIDN:AAD20917.1; GSPDB:GN00139

C:Genetics:

A:Gene: AT2g20780

A:Map position: 2

C:Superfamily: glucose transport protein

Query Match 46.0%; Score 1176; DB 2; Length 547;

Best Local Similarity 45.7%; Pred. No. 2.8e-78;

Matches 241; Conservative 96; Mismatches 138; Indels 52; Gaps 6;

QY 8 KAVEPRKGNVYASICAILASNASVILGY-----DIGVMSGAA 46

DB 42 REAEARNRTRKYVACAFASLNNVLLGYGRFYLNRLILLVFDLQKDVGMVSGAV 101

QY 47 MYIKDLNITDVQLEILIGILSYLSFGSFAGARTSDRIGRLTVVFAVIFVGSLLMG 106

Db 102 LFIQDLKITEVQTEVLIGLSLSIISLFGSLAGRTSDSGRKTWMAALVFTQCAVMA 161

QY 107 FAINTYGLMAGRFVAGVGVGGMIAPIYTAISPAASRGFLTTPPEVFINIGILLGVL 166

Db 162 VAPSEVLMIGTLTLAGIGLGMIAPIYTAISPTVARGFTSPPEIFINIGILLGV 221

QY 167 NFAFARPLHLGWRVMAITCAVPSGLALLVFCMPESPRLVLKRLADARAVLEKTSAT 226

Db 222 NYAFGLSVHISWRMLAVILPSVIFGALCVIPESPRLVMKGRVDSAREVLKMTNER 281

QY 227 PBEAERLADIKAAGIPKGLDGVVTPGKQGGELQVWKLLSTPAPVRRILLASV 286

Db 282 DDEABERLAEIOLAA-----AHTEGSEDRPVWREL--LSPSPVVRKMLVGF 326

QY 287 GLHFFQAGSDSVVOYSARLFKSAGITDDNKLGVTCVAGVTKTFFILVATFLDRAGR 346

Db 327 GQCFOQITIGIDATVYSPSEILKEAGIQDETLLAATVAVGVTKVFLFATFLDSVGR 386

QY 347 RPLLLISTGGMIVSLICLGGSTVAGHHPDPTKVAVAVLCTASTLSYTAFFSIGLPITG 406

Db 387 KPLLYVSTIGMTLCLPCLSTFLTLG-----QGTGLITALLFVGCNVAFFSIGMPVCM 441

QY 407 VYTSEIFPPLQVAGVAVGNRVTSVAVISMTLSKAITIGSGSEFSLYSGIAVAVVF 466

Db 442 VITSEIFPPLRAQASALGAVGNRVCSGLVAVSVLSKAITVGGTFFVFLSVLSVIF 501

QY 467 FTFCLPTEGRTRILEMGLF-----GMPDTGMAEEAEDAAKEK 505

Db 502 VYVLVETSGKSLQELMFGGLRKGDEVELG---DAERLVRKEQ 545

RESULT 7

E70070

Metabolite transport protein homolog ywEG - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000

C:Accession: E70070

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertei, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallet, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, R.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: E70070

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-457 <KUN>

A:Cross-references: GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15600.1; PID:G2636109

A:Experimental source: strain 168

C:Genetics:

C:Superfamily: glucose transport protein

Query Match 26.1%; Score 668; DB 2; Length 457;

Best Local Similarity 32.4%; Pred.No. 2.9e-41;

Matches 147; Conservative 99; Mismatches 176; Indels 32; Gaps 7;

QY 29 SMASVILGYDVGVMGAAMVIRKDLNITDVLQELIGLSLSYLFSGSPAGATSDRIGRR 88

Db 14 ALGGALYGYDVTGVISGAILFMKELGNLAFTEGLVSSLLVAILGSGAAGKLTDRFGK 73

QY 89 LTVVFAVFFVCSLLMGFAVNVGMAGRFVAGVGVGGMIAPIYTAISPAASRGFL 148

Db 74 KATMAAALLFCIGGLGVALAPNTGMVLFRIILGLAVGTSTTIVPIYLSLAPKHKRGAL 133

QY 149 TTPPEVFINIGILLGYSNFAFARLPLHLGWRVMAIGAIVPSGLLALLVFCMPESPRLV 208

Db 134 SSINQLMITVIGLLSYVNVYIFADAB--AWRWMLGLAAVPSLLILLIGLFWPESPRWLF 190

QY 209 LKRLADARAVLEKTSATPEEAAERLADIAAAGIPKGLDGVVTPGKQGGELQVWK 268

Db 191 TNGEESKAKILEKLRGT-KDIDQETHDIKEAE-----KODEGGLKEIF- 233

QY 269 KLILSTPAPVRRILLASVGLHFFQAGSDSVVOYSARLFKSAGITDDNKLGLVTCVAGV 328

Db 234 -----DPWVRPALIAGLGLAFLOQFIGNTIIYAPKFTTNVFGNSASILG-TVGIGT 286

QY 329 TKTFFILVATFLDRAGRPPLLISITGGMIVSLICLGGSTVAGHHPDPTKVAVAVLCTA 388

Db 287 VNVLMTLVAIKIDKIGRKPPLLFGNAGMVISLIVLALVNLFFNNTP--AASWTTVICLG 344

QY 389 STLSYTAFFSIGLGPITGVYTSIFLPQVRLGFAVGAVSNRVTSVAVISMTLSLSKAIT 448

Db 345 ---VFIVFAVSGPVVWVWMLPELFFPHVRGIGTGVTMLHVGTLIVSITPILMEALG 401

QY 449 IGGSEFFLYSGIAAVAVWVFFFTCLPTEGRTRILEM 482

Db 402 ISYLFLLIYAAIGTMAFLFVRFKVTETKGRSLBEI 435

RESULT 8

D86426

hypothetical protein F12P21.2 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002

C:Accession: D86426

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-580 <STO>

A:Cross-references: GB:AE005172; NID:gl1120774; PIDN:AAG30955.1; GSPDB:GN00141

C:Genetics:

C:Map position: 1

C:Superfamily: glucose transport protein

Query Match 24.6%; Score 629.5; DB 2; Length 580;

Best Local Similarity 29.5%; Pred.No. 2.5e-38;

Matches 161; Conservative 95; Mismatches 193; Indels 97; Gaps 10;

QY 20 YASICALASMASVILGYDVGVMGAAMVIRKDLNITDVLQELIGLSLSYLFSGSF 76

Db 27 YVLELAFSAGIGLLFGYDVGIVSGALLIYRDDFKSVDRNTWLQEMVMAVAGAI 86

QY 77 AGARTSDRIGRLTVVPAAVIFPVGSLMGFAVNVGMAGRFVAGVGVGGMIAPIYVT 136

Db 87 ICGWANDKLGRSAILMADEFLILGAIIMAAANPSSLVVRVGVGVGVMASNTAPLYI 146

QY 137 AEISPAASRGFLTTPPEVFINIGILLGYSNFAFARLPLHLG-WRVMLAIGVPSGLAL 195

Db 147 SEASPAKIRALYSTNGFLITGQGFSLYNLAFTDV---TGTWRWMLGAGIPALLQV 203

QY 196 LVFCMPESPRLVLMKRLADARAVLEKTSATPEEAAERLADIKAAGIPKGLDGVVTP 255

Db 204 LMTPLPSPRLVLRKGRBEAKAILRIYSA-EDVEQETRAK-----DSVETE 251

QY 256 GREQGGELQVWKKLILSPTPAVRRILLAVGLHFFQOASGDSVVQYSARLFKSGAGITD 315  
DB 252 ILEGSSEKINMIKCKAKT--VRRGLIAGVGLVQFQVGVNTVMYSPFTVQLAGPAS 309  
QY 316 DNKLLGVTCVAVGVTKFTFFILVATFLDRAGRPRLILISTGGMIVSLICLGSGLTVAGHH- 374  
DB 310 NRTALLSLVTAAGLNAFGSIIISFYIDR:GRKKLLIISLFGVILSGILTVFYEAATHA 369  
QY 375 -----PDKTKV-----AWAVALCI-ASTLS----- 392  
DB 370 PAISSLETOFNNISCPDYKSNMNTNWDNCTLKASPSGCYCSPIGKEHPGACWISD 429  
QY 393 -----YIAFFSIGLPGITGVYTSBIFLQVRA 419  
DB 430 DSVKOLCHNENRLWYTRGCPNSFGFWALLGLGLYIIFPSPGMGTVPEWINSIYELFRFG 489  
QY 420 LGFAVGVASNRVTSAVISMTFLSLKAITIGSPFLYSGIAAVANVFFTCLPETRGRTL 479  
DB 490 ICGGTAATNWSLNLIVAOFSLSLLEAIGTSMTFLIFGVISVIALLFVWVCVPETKGPMP 549  
QY 480 BEMGKL 485  
DB 550 EIEIKM 555

RESULT 9  
D70073  
metabolite transport protein homolog yxoc - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: D70073  
R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasunoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: D70073  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-461 <KUN>  
A:Cross-references: GB:Z99124; GB:AL009126; NID:92636442; PIDN:CAB16017.1; PID:g2636527  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yxoc  
C:Superfamily: glucose transport protein

Query Match 24.4%; Score 624.5; DB 2; Length 461;  
Best Local Similarity 33.2%; Pred. No. 4.4e-38;  
Matches 157; Conservative 85; Mismatches 190; Indels 41; Gaps 12;  
QY 14 KKGNYKASICAILLASMASVILGVDIGVMSGAAMVYKLDNI-----TDVOL 60  
DB 2 KKDTRKY--MIYFFGGLGLLYGYDTGVISGALLFINNDPLTLTLEGLVMSLLIGAIF 59  
QY 74 GSPAGARTSDRIGRLTVFAAVTFFVGLSLMGFAVNYGMLMAGRFVAGVGVGGMIAP 133  
DB 60 GSALSGTCSDRWGRKRVVFLSIIFIICALACAPSQIGMLIASRVILGLAVGSGTALVP 119  
QY 134 VYTAISGPAASRGFLTTFPEVFINIGILLGLVLSNFARLPLHLGWRVMLAIGVPSGLL 193  
DB 120 VYLSEAPTIRGLTGTWNMLMIVTGILLAYINYLFT--PFE-AMRWVGLAARVAVLL 176

QY 194 ALLVFCMPESPRWLVKRLADARAVLEKTSATSEEAERLADIKAAAGIPKGLDGDVVT 253  
DB 177 LIGAFMPESPRWLVRKGSBEAARMNITH-DPKDIEMELAENK----- 220  
QY 254 VPGKEQGGELQVWKKLILSPTPA--VRRILLAVGLHFFQOASGDSVVQYSARLFKSA 311  
DB 221 -----QGEAEK--KEITLGLVKAKWRPMLLIGVGLAIQQAVGINTVIYATITKA 272  
QY 312 GITDNKLLGVTCVAVGVTKFTFFILVATFLDRAGRPRLILISTGGMIVSLICLGSGLTVA 371  
DB 273 GLGTSASALG-TMGIGILNVIMCITAMILIDRVGRKKLLIWSGVGITLSLAALSGVLLT 331  
QY 372 GHHPDPTKVAVALCIASLTLSYIAFFSIGLPGITGVYTSBIFLQVRA--ALGFAVGVASN 429  
DB 332 GLSAST--AMTVVFLG--VYIVFYQATWGPVWVLMPELFFSKARGAATGFTTLVLS- 385  
QY 430 RVTSASVMTFLSLKAITIGSPFLYSGIAAVANVFFTCLPETRGRTLEEM 482  
DB 386 -AANLIVSLVFLMLSANGIAWFMVFSVLCILSFFFAFYVMVETKGSLEEI 437

## RESULT 10

G84864

Probable membrane transporter [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: G84864

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.J.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: G84864

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-521 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g43330

A:Map position: 2

C:Superfamily: glucose transport protein

Query Match 23.7%; Score 606; DB 2; Length 521;

Best Local Similarity 29.9%; Pred. No. 1.1e-36;

Matches 153; Conservative 93; Mismatches 178; Indels 88; Gaps 10;

QY 16 GNKXASICAILLASMASVILGVDIGVMSGAAMVYKLDNI-----TDVOL 60

DB 27 GN-SYILGTVTAGIGLLFGYDTGVISGALLYIKDDFEVVKQSSFLOVYVNSFTSSKL 85

QY 61 EILIGILSLYLFSGPAGARTSDRIGRLTVFAAVTFFVGLSLMGFAVNYGMLMAGRFV 120

DB 86 ETIVSMALVGMIGAAAGGWINDIYGRKATLFDVVPFAAGAVMAAADPPVILISGRLL 145

QY 121 AGVGUGYGMIAPVYTAISPAASRGFLTTFPEVFINTIGILGLYSNFAPARLPHLG-W 179

DB 146 VGLGVGVASVTAPVIAEASPSVGRGLVSTNVLMTITGGQFLSYLVNSAFTQVP---GTW 202

QY 180 RVMLAIGAVPSGLLALLVFCMPESPRWLVKRLADARAVLEK-----SAT 226

DB 203 RWMGLGVGPVAVIQFILMFEMPESRWLFMKRKAEQVLAITYDIRLEIDEHLGAA 262

QY 227 PEEAERLADIKAAAGIPKGLDGDVTVVPKGGQGGELQVWKKLILSPTPAVRILISAV 286

DB 263 EEEKQRKRTV-----GYLDVFR-----SKELRLAFLAGA 292

QY 287 GLHFFQOASGDSVVQYSARLFKSAIGITDNDKLLGVTCVAVGVTKFTFFILVATFLDRAGR 346

DB 293 GLQACQFGTGVNTVMYSPFTVQMGAFHSNQALFSLIVAAAMNAGTGVWGYIFDHCGR 352

QY 347 RELLILISTGGMIVSLICLG-----SGLTVAGHHPDPTKVAWAVALCIASLTLSVY 394

353 KKLALSLFLGVIIISLLSVSFFKQSTSDGGL-----YGLAVLGLA---LYI 399  
Db  
395 AFFSIGLGPITGVYTSIFLPLQVRLGFAVGASNRVTSIAVISMFTLSLSKAITIGGSFF 454  
QY  
400 VFFAPGPGVPWTVNSIYFQQVGGICGGSATWNISLIVAQTLTTAAGTGMFL 459  
Db  
455 LYSIGIAAVAVVFPFTCLPPTGRITLBMGKLF 486  
QY  
460 ILAGIYLAIVFIVFVFPETQGLTFSEVEQIW 491  
Db  
RESULT 11  
F71431  
hypotheical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
A:Variety: colombia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: F71431  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk  
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel  
avanagh, T.; Hempel, S.; Kotter, P.; Enrian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Monfort, A.; Pons, A.; Puigdomenech  
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; An  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal  
A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: F71431  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-582 <BEV>  
A:Cross-references: GB:297341; NID:G2244991; PID:G2245004  
C:Genetics:  
A:Map position: 4COP9-4G3845  
C:Superfamily: glucose transport protein

Query Match 23.3%; Score 597; DB 2; Length 582;  
Best Local Similarity 27.8%; Pred. No. 5.9e-36;  
Matches 160; Conservative 99; Mismatches 209; Indels 108; Gaps 11;  
QY 1 MASDELAKAVEP-----RKKNVYKVASICAILASMASVILGVDIGVMSGAAMYIKDNLN 54  
Db 1 MVEGGIAKADKTECTWKTTPYIMRLALSAGIGLLFGYDGTGVSIGALLFIKEDFD 60  
QY 55 ITDVOELIGILSLY---SLFSGFAGARTSDRIGRLTWFAAVIFFVGSLLGFAVNY 111  
Db 61 EVDKKTWLTQSTIVSMVAGAVGAAGVGGWINDKFGGRMSILIAVLFLGIAVMAFAP 120  
QY 112 GMLMAGRFVAGVGVGGMIAPIVYTAIEISPAASRGFLTTFPEVFNIGILLYSNFAPA 171  
Db 121 WLIIVGRIFGVGVGMAWMTSPLYISEASPARIRGALVSTNGLLITGGOFFSYLINLAFV 180  
QY 172 RLPLHLG-WRMLAIGAVPGLLALVFCVPSRMLVLKGRADARAVLEKTSATPEEA 230  
Db 181 HTP---GTWMLIGVAGVPAIVQVLMLSLTPSPRMLYKRDIAESAILERIYPADAVE 237  
QY 231 AERLADIAKAAAGIPK---GIDGQWTVPGKEGGGSLQWKKLIILSPFAVRRIILLSAVG 287  
Db 238 AEMEA-LKLSVEAKEADEAIGDSFSAKLKAGF-----NPPVRRGLAAGIT 283  
QY 288 LHFFQQASGSDSVVQVSARLFKSAGITDDNKLGVTCVAVGVTWTFILVAFPLDRAGR 347  
Db 284 VQVAQQPVGINTWYISPSIVQAGYASNKTAUMLSLITSGNALGSIWMMFVDYGR 343  
QY 348 PLLIISGGMIVSLICLGSGLTVAGH-----PDKVA----- 380  
Db 344 KLMISWFGIILACILITAVFSQAALHAPKIDAFESRIFAPNATCSAYAPLAENAPSR 403  
QY 381 WAVALCIAS-----TLS----- 392  
Db 404 WNCWKILRSRCEGCASGVQYAPGACVILSDDKMKTATCSRGRITFKDQCPKFGFLAIVF 463  
QY 393 ---YIAFFSIGLGPITGVYTSIFLPLQVRLGFAVGASNRVTSIAVISMFTLSLSKAITI 449

464 LGLYIVVYAPGMCVTPWIVNSEIYPLRYRGLGGIAAASNWVSNLVSESFLSLTHALGS 523  
QY 450 GGSFFLYSGIAAVAVVFPFTCLPPTGRITLBMGKLF 485  
Db 524 SGIFLLFAGFSTIGLFFIWLVPETKGLQFEEVEKL 559  
Db  
RESULT 12  
T51485  
sugar transporter-like protein - Arabidopsis thaliana  
N:Alternate names: protein T21H19\_70  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: T51485  
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew  
submitted to the Protein Sequence Database, August 2000  
A:Reference number: Z25394  
A:Accession: T51485  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-560 <SAT>  
A:Cross-references: EMBL:AL391148  
A:Experimental source: cultivar Columbia; BAC clone T21H19  
C:Genetics:  
A:Map position: 5  
A:Insertions: 86/1; 124/2; 161/1; 205/2; 238/3; 276/2; 307/3; 359/1; 408/1; 432/3; 462/2; 50  
A:Notes: T21H19\_70  
C:Superfamily: Glucose transport protein

Query Match 23.3%; Score 595; DB 2; Length 560;  
Best Local Similarity 30.7%; Pred. No. 7.9e-36;  
Matches 146; Conservative 91; Mismatches 186; Indels 52; Gaps 9;  
QY 27 LASMASVILGVDIGVMSGAAMYIKDNLNITDVLQILIGILSLYSLE----- 73  
Db 111 VACIGAILFGYHLGVNGALEYLAIDLGI--AENTVLOGKYMHIHFFPPVNGWIVSLL 168  
QY 74 ---GSPAGARTSDRIGRLTWFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVY 128  
Db 169 AGATVGSFTGGALADKFCRTRTFQDAIPALGAFLCATAQSVQTMIVGRLLAGIGIS 228  
QY 129 GMTAPVYTAIEISPAASRGFLTTFPEVFNIGILLYSNFAPARLPLHLGWRVMLAICAV 188  
Db 229 SAIVPIYIIEISPTIRGALGSVNLFCIGILALIAIGLPLAANP--LWRTWFGVAVI 286  
QY 189 PGGLLALLVFCVPSPRMLVLKGRADARAVLEKTSATPEEAERLADIKAAGIPKGLD 248  
Db 287 PSVLLAIGMAFSPSPRMLVQGVKUSEAKAI-KTYGKERVWELVRDLASG----- 338  
QY 249 GDVTVPGKEGGGLOV-WKKLIILSPFAVRRIILLSAVGLHFFQQASGSDSVVQVSARL 307  
Db 339 -----QSSPEAGNFDLSS---RYWVVSVCAALFLFQQLAGINAVVYKTSV 385  
QY 308 FKSAGITDDNKLGVTCVAVGVTWTFILVATFLDRAGRPLLIISTGGMIVSLICIGSG 367  
Db 386 FRSAIGIQSD---VAASALVGSNVFETAVASSLMDKMKRKSLLTSTFGMALSLLSLS 442  
QY 368 LTVAGHHPDTKVAMAVALCIATSLSYIAFFSIGLGPITGVYTSIFLPLQVRLGFAVGA 427  
Db 443 FTW-----KALAAISGTLAVGTVLVLSFSLGAGVPALLLPFASIRINAKAVALSIG 497  
QY 428 SNRVTSAVISMFTLSLSKAITIGGSFFLYSGIAAVAVVFPFTCLPPTGRITLBM 482  
Db 498 MHWISNFVIGLYFLSVWTKFGTSSVYLGFAGVGVCLVAVLYIAGNVVETKGRSLEBI 552  
Db  
RESULT 13  
D96539  
hypotheical protein F1413.9 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: D96539



[illegible]

A;Status: nucleic acid sequence not shown; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-472 <MAI>  
P;Maiden, M.C.J.; Jones-Mortimer, M.C.; Henderson, P.J.F.  
J. Biol. Chem. 263, 8003-8010, 1988  
A;Title: The cloning, DNA sequence, and overexpression of the gene araE coding for arabinose efflux pump  
A;Reference number: A28075; MUID:88228015; PMID:2836407  
A;Accession: A28075  
A;Molecule type: DNA  
A;Residues: 1-472 <MA2>  
A;Cross-references: GB:J03732; NID:g145320; PIDN:AAA23459.1; PID:g145321  
R;Stoner, C.; Schleif, R.  
J. Mol. Biol. 171, 369-381, 1983  
A;Title: The araE low affinity L-arabinose transport promoter. Cloning, sequence, transcription, and regulation  
A;Reference number: I40996; MUID:84114868; PMID:6319708  
A;Accession: I40996  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-25 'Y' 349, 'R' <RES>  
A;Cross-references: EMBL:X00272; NID:g40940; PIDN:CAA25075.1; PID:g40941  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A;Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of *Escherichia coli* K-12.  
A;Reference number: A6720; MUID:97426617; PMID:9278503  
A;Accession: B65067  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-472 <BAT>  
A;Cross-references: GB:AE000368; GB:U00096; NID:g2367165; PIDN:AAC75880.1; PID:g1789207;  
A;Experimental source: strain K-12, substrain MG1655  
C;Genetics:  
A;Gene: araE  
A;Map position: 61 min  
C;Superfamily: glucose transport protein  
C;Keywords: arabinose transport; intramolecular oxidoreductase; isomerase; membrane protein

Query Match	22.9%	Score 586.5	DB 2	Length 472
Best Local Similarity	29.9%	Pred. No. 2.7e-35		
Matches 147	Conservative 104	Mismatches 197	Indels 43	Gaps 9
QY	1	MASDELAKEVPRKGNKVYASI	-CAILASMASVILGYDGYVMGGAAMVYKIKDLNITDVQ	59
Db	1	MVTINTESALTPSLRDLTRRMNMFVSAVAAGCLFLGLDGVJAGALPFTIDHFVITSRL		60
QY	60	LEILIGILSYLSFGSFAGARTSDRIGRLTLVVFAAVIFVGSLLMGPAVNYGMLMAGRF	119	
Db	61	QEWVSSMMLGAAGALFNGLSWFLGRGRKYSLMAGAILFVLGSIAGFAFATSVEMLIAARV	120	
QY	120	VAGVGVGYGMTAPVYTAISPAASRGFTTTPPEVINTGILLYLSNFAFARLPJHLG-	178	
Db	121	VIGIAVGIASYTAPLYLSEWSENVRGMSVYQLMVTTLGIVLAFUSDFAFS----	176	
QY	179	WRVMLAIGAVPGLLALLVFCMPESPRWLVLKRLADARAVLEKTSATPEAAERLADIK	238	
Db	177	WRAMLGLVALPAVLLIILVVFLPNSPRWLAEGRHTEAEVLRMLRDTSEKAREELNEIR	236	
QY	239	AAAGIPKGLDGDVTVTPGKEGGELQVWKLIILSTPPAVRILLSAVGLHHFFQOAGSD	298	
Db	237	ESLKI-----KQSG-----WALFKINRN--VRAVFLGMLLQAWQQQTGMN	275	
QY	299	SVVQYSARLFKSAGITDDNKLGLVTOAVGVTXTFFILVATFLDLDRAGREPLLISITGMI	358	
Db	276	IIMYAPRIKFMAGETTTTQQQMIATLVGLTTFMFATFAVFTVDKAGRXPALKIGSVMA	335	
QY	359	VSILCLG-----SGLTVAGHHPTKVAAVAVALCIATLSLYAFPSIGLGPITGVYTSR	411	
Db	336	LGTVLVGLCMLQFDNGTAGS-----LSW-----LSVGTMMCIAGVMSAAPVVWILCSR	386	
QY	412	IFPLQVRALGFVAVGNASNRVTSVATSMFTLSLSUKAITIGGSPFLYSGLIAAVAWVFFTC	471	
Db	387	IQLKCRDPIGTCTSTTTTNNVWSNMIIIGATFLTLLDSIGAAGTFLWLTALMIAFVIGITFW	446	

RESULT 14

RESULTS  
B26430  
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain Y-13)

C:Species: *Escherichia coli*

C:\Data\1988\05-Oct-1988\sequence revision 05-Oct-1988 #text change 01-Mar-2002

C;Accession: B26430; A28075; I40996; B65067

R;Maiden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.F.

Nature 325, 641-643, 1987

A;Title: Mammalian and bacterial sugar transport proteins are homologous.

A;Reference number: A93389; MUID:87115869; PMID:3543693

A;Accession: B26430

QY 472 PETRGRTELEM 482  
DB 447 PETKNVTLEHI 457

RESULT 15  
B91091  
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain O157:H7, substrain RIMD 05  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 31-Dec-2001  
C:Accession: B91091  
R:Havashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: B91091  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-472 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA37121.1; PID:g13363170; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECs3698  
C:Superfamily: glucose transport protein  
C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 22.9%; Score 586.5; DB 2; Length 472;  
Best Local Similarity 29.9%; Pred. No. 2.7e-35;  
Matches 147; Conservative 104; Mismatches 197; Indels 43; Gaps 9;

QY 1 MASDELAKAVEPRKGNVYASICAILASMASVILGYDVGMSGAAMYIKKDLNITDVQ 59  
DB 1 MVTINTESALTFRSLRDTFRMMNFVSAAGVAGLLFDIGDIAGALPFTIDHFEVLTSL 60

QY 60 LEILGILSLYSLFSGFAGARTSDRIGRLTVVFAVFFVGSLLMGFAVNYGMLMAGRF 119  
DB 61 QEWVSSMMLGAIGALFNGWLSFRLGRKYSLMAGAILFVLGSGSFAFATSVEMLIAARV 120

QY 120 VAGVGVGCGMIAPVYTABISPAASRGFLTTPPEVFINIGILLGYLSNFAPARLPHLG- 178  
DB 121 VLGIAGIASYTAPLYLSEMASENVRGKMSMYQLMVTILGIVLAFSLDTAFS---YSGN 176

QY 179 WRVMAIGAIVPSGLLALLVFCMPESPRLVLKRLADAPAVLEKTSATPEAAERLADIK 238  
DB 177 WRAMLGVLAPVLLIILVFLPNSPRWLAKEGRHIEAEVLRMLRDTSEKAREELNEIR 236

QY 239 AAAGIPKGLDGDVTVTPGKEQGGLQVWKKLILSPTPAVRILLISAVGLHFFQQAQSD 298  
DB 237 ESLK-----KQSG-----WALFKINEN--VRAVFLGMLQAMQOFTGMN 275

QY 299 SVVQYSARLFKAGITDDNKLGLVTCAGVVTFFILVATFLLDRAGRPLLIIISTGGM 358  
DB 276 IIMYAPRIFKQAGFTTTEQOMIATLVGLTFMFATFIAVFTVDKGRKPKLKIGFSYMA 335

QY 359 VSLICIG-----SLTVAGHHDPDKVAVAVALCIASLTLSVIAFPFISIGLGPITGVYTS 411  
DB 336 LGTLVLGYCLMQPDNQTASSG-----LSW---LSVGMWMCIAGYMSAAPVWVILCSE 386

QY 412 IFPLQVRLGFAVGASNRVTSVAVISMTFLSLSKAITIGGSFFLYSGIAAANAVWFFFTCL 471  
DB 387 IQPLKCRDFGITCSTTTNVSNMIIAGATFLLTDSIGAAGTFWLYTALNIAFVGITFWLI 446

QY 472 PETRGRTELEM 482  
DB 447 PETKNVTLEHI 457

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:31:52 ; Search time 17 seconds  
(without alignments)  
1571.294 Million cell updates/sec

Title: US-10-051-902A-20

Perfect score: 2559

Sequence: 1 MASDELAKAVEPRKGNVYK.....AEEAEDAAKKEKVVLPSSK 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642.5	25.1	629	1 MYCT_HUMAN	O96ge2 homo sapien
2	624.5	24.4	461	1 CSBC_BACSU	P46333 bacillus su
3	591.5	23.1	457	1 XVT_LACBR	O52733 lactobacill
4	586.5	22.9	472	1 ARAE_ECOLI	P09830 escherichia
5	586	22.9	557	1 ITR2_SCHPO	P87110 schizosacch
6	579	22.6	472	1 ARAE_KLEOX	P45598 klebsiella
7	568.5	22.2	482	1 YFIG_BACSU	P54723 bacillus su
8	565	22.1	584	1 ITR1_YEAST	P30605 saccharomyc
9	563	22.0	612	1 ITR2_YEAST	P30605 saccharomyc
10	547	21.4	464	1 GALP_ECOLI	P37021 escherichia
11	546	21.3	575	1 ITR1_SCHPO	Q10286 schizosacch
12	544	21.3	491	1 XYLE_ECOLI	P09098 escherichia
13	538	21.0	464	1 ARAE_BACSU	P96710 bacillus su
14	526	20.6	468	1 GLCP_SYNY3	P15729 synechocyst
15	522	20.4	534	1 HUP3_CHLKE	Q39525 chlorella k
16	520	20.3	547	1 GTR1_LEIDO	Q01440 leishmania
17	515.5	20.1	477	1 GTR8_HUMAN	Q9ny64 homo sapien
18	515.5	20.1	522	1 STPI_ARAPH	P23586 arabidopsis
19	509.5	19.9	477	1 GTR8_MOUSE	Q9jif3 mus musculu
20	505.5	19.8	494	1 GTR3_SHEEP	P47843 ovis aries
21	505.5	19.8	495	1 GTR3_CANFA	P47842 canis fami
22	504.5	19.7	523	1 STC_RICCO	Q41144 ricinus com
23	502.5	19.6	534	1 HUP1_CHLKE	P15686 chlorella k
24	501.5	19.6	494	1 GTR3_BOVIN	P36352 bos taurus
25	498	19.5	478	1 GTR8_RAT	Q9j7jz rattus norv
26	489.5	19.1	493	1 GTR3_RAT	Q07647 rattus norv
27	489.5	19.1	510	1 HEX6_RICCO	Q07423 ricinus com
28	486	19.0	490	1 GTR1_CHICK	P46896 gallus gall
29	480.5	18.8	524	1 GTR2_HUMAN	P11168 homo sapien
30	476.5	18.6	523	1 GTR2_MOUSE	P14248 mus musculu
31	474	18.5	492	1 GTR1_RABIT	P13355 oryctolagus
32	473	18.5	492	1 GTR1_RAT	P11167 rattus norv
33	471	18.4	492	1 GTR1_MOUSE	P17809 mus musculu

34	470.5	18.4	493	1 GTR3_MOUSE	P32037 mus musculu
35	468	18.3	492	1 GTR1_HUMAN	P11166 homo sapien
36	464	18.1	496	1 GTR3_CHICK	P28568 gallus gall
37	463	18.1	496	1 GTR3_HUMAN	P11163 homo sapien
38	462	18.1	592	1 HXT5_YEAST	P38695 saccharomyc
39	461	18.0	492	1 GTR1_BOVIN	P27674 bos taurus
40	459.5	18.0	541	1 GTR1_HUMAN	O35528 homo sapien
41	459	17.9	540	1 HUP2_CHLKE	Q39524 chlorella k
42	457.5	17.9	522	1 GTR3_RAT	P12336 rattus norv
43	451.5	17.6	522	1 STA_RICCO	Q10710 ricinus com
44	447.5	17.5	509	1 GTR4_MOUSE	P14142 mus musculu
45	447.5	17.5	533	1 QUTD_EMENI	P15325 emericeella

#### ALIGNMENTS

##### RESULT 1

ID	MYCT_HUMAN	STANDARD;	PRT;	629 AA.
AC	O96GE2;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Proton myo-inositol co-transporter (Hmit).			
GN	SLC2A13.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21391710; PubMed=11500374;			
RA	Udrys M., Ibberson M., Horisberger J.-D., Chatton J.-Y.,			
RA	Riederer B.M., Thorens B.,			
RT	"Identification of a mammalian H(+)-myo-inositol symporter expressed			
RT	predominantly in the brain."			
RL	EMBO J. 20:4467-4477(2001).			
CC	-!- FUNCTION: H(+)-myo-inositol co-transporter. Can also transport			
CC	related stereoisomers.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- TISSUE SPECIFICITY: Predominantly expressed in the brain.			
CC	-!- PTM: Glycosylated.			
CC	-!- SIMILARITY: Belongs to the sugar transporter family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL: AJ315644; CAC51116.1; -			
DR	Genew; HGNC:15956; SLC2A13.			
DR	InterPro; IPR007114; MFS.			
DR	InterPro; IPR005828; Sub transporter.			
DR	InterPro; IPR005829; Sug transporter.			
DR	InterPro; IPR003663; Sugar transpt.			
DR	Pfam; PF00083; sugar tr; 1-			
DR	PRINTS; PR00171; SUGTRNSPORT.			
DR	TIGRfam; TIGR00879; SP; 1.			
DR	PROSITE; PS00850; MFS; 1.			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
KW	Transmembrane; Sugar transport; Glycoprotein.			
FT	DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 58 78 1 (POTENTIAL).			
FT	DOMAIN 79 106 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 107 127 2 (POTENTIAL).			
FT	DOMAIN 128 129 CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 130 150 3 (POTENTIAL).			
FT	DOMAIN 151 159 EXTRACELLULAR (POTENTIAL).			

FT	TRANSMEM	160	180	4 (POTENTIAL).
FT	DOMAIN	181	193	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	194	214	5 (POTENTIAL).
FT	DOMAIN	215	220	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	221	241	6 (POTENTIAL).
FT	DOMAIN	242	305	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	306	326	7 (POTENTIAL).
FT	DOMAIN	327	344	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	345	365	8 (POTENTIAL).
FT	DOMAIN	366	374	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	375	395	9 (POTENTIAL).
FT	DOMAIN	396	489	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	490	510	10 (POTENTIAL).
FT	DOMAIN	511	530	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	531	551	11 (POTENTIAL).
FT	DOMAIN	552	554	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	555	575	12 (POTENTIAL).
FT	DOMAIN	576	629	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	414	414	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	439	439	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	466	466	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SEQUENCE	629 AA;	68203 MW;	8AD79C56C790814E CRC64;
Query Match		25.1%;	Score 642.5;	DB 1; Length 629;
Best Local Similarity		28.9%;	Pred. No. 2e-35;	
Matches 158;		Conservative 88;	Mismatches 205;	Indels 95; Gaps 3;
QY	20 YASICAILASMASVILGYDIGWSGAAMYTKDNLITDVLQSEILLIGILSLYSLFGSPAGA	79		
DB	61 FVYVAVFSLGGLFGYDYGWSGAMLLKRLQSLDALQBELSLVSTVGAAGAVALAGG	120		
QY	80 RTSDRIGRLTVFAAVIYFFVGLLMGFANVYGMAGRFVAGVGVGYGGMIAPIVYTAEI	139		
DB	121 ALNGVFGRAAILLASALFTAGSAVLAANAANKETLLAGRLVGLGIGIASMTVPVYIAEV	180		
QY	140 SPAASRGFLITTFPEVFNIGILLGYSNFAFARLPLHLGWRVLAIGA VPSGLLALLVFC	199		
DB	181 SPPNLRGLVINTLFTITGGQFFASVVDGAFSLQKD-GWRYMLGLAXVPAVIQFFGFLF	239		
QY	200 MPESPRMLVLKRLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGDVVTVPGKEQ	259		
DB	240 LPESPRMLIQGQTKARRILSQMRGN-QTIDSEYDSIK	287		
QY	260 GGGEQLQWKKLILSPTPAVRILLASVGLHFFQASGSDSVVOYSARLPKFSAGITDDNKL	319		
DB	288 GSAGPVICRMLSYPPT---RREALVGCGLQMFQQLSGINTIMYVSATILQMSGVEDRLA	344		
QY	320 LGVTCAGVTKTFFILVATFLDRAGRRPLLSTGGMVSLVICLSGLTVAGH-----	373		
DB	345 IWLASVTAFTHFIETLVGLVMLEKVGKRKLTFCGLAGTTVALIILALGPVLSAQVSPRIT	404		
QY	374 -----HPD-----	376		
DB	405 FKPIAPSQNATCTYSYCNECMLDPCGFCXKMNKSTVIDSSCPVFNKASTNEAAWGRC	464		
QY	377 -----TK-----VANAVALC-----IASTLSVIAFFSGTGPITGVYTSFFLPQVRL	420		
DB	465 ENETFKTEDIFWYNFCPTPYSTALLGLILVLVFPAGMGMPMTWVNSEIPLUWAKST	524		
QY	421 GFAGVAGSNRVTSVAISMFTLSLKAITGGSPFLYSGIAAANAWVFFFCLPETERGRILE	480		
DB	525 GNACSSGNNWTFNVLSLTFHTAEYLYTYYGAFVLAGFAAGVGLLIYVGLCPETGKKKLE	584		
QY	481 EMGKLF	486		
DB	585 EIESLF	590		
RESULT 2				
ID	CSBC_BACSU	STANDARD;	PRT;	461 AA.
AC	P46333; Q32289;			
DT	01-NOV-1995 (Rel. 32, Created)			

```

10-OCT-2003 (Rel. 42, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable metabolite transport protein csbC.
CSBC OR S592BR OR BSU39810.
Bacillus subtilis.
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_taxid=1423;
[1]
SEQUENCE FROM N.A.
STRAIN=168 / BGSC1A1;
MEDLINE=96093926; PubMed=7584049;
Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
"Cloning and sequencing of a 36-kb region of the Bacillus subtilis
genome between the gnt and iol operons.";
DNA Res. 2:61-69(1995).
[2]
REVISIONS.
Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
STRAIN=168;
MEDLINE=98044033; PubMed=9384377;
Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.-K., Codani J.J., Conterton I.F., Cummings N.J., Daniel R.A.,
Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
Entian K.D., Ewington J., Fabret C., Ferrazi E., Foulger D.,
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
Joris B., Karamata D., Kashiwa Y., Klammer-Blanchard M., Klein C.,
Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
Kurita K., Lapidus A., Lardinio S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
Presecan E., Fujita P., Furnelle B., Rapoport G., Rey M., Reynolds S.,
Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Sekiguchi J., Sekowska A., Sertor S.J., Serror P., Shin B.S., Solido B.,
Sorkin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
Tosatti V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
"The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis";
Nature 390:249-256(1997).
[4]
CHARACTERIZATION.
MEDLINE=99303315; PubMed=10376822;
Akbar S., Lee S.Y., Boylan S.A., Price C.W.;
"Two genes from Bacillus subtilis under the sole control of the
general stress transcription factor sigmaB.";
Microbiology 145:1069-1078(1999).
CC -!- FUNCTION: Could serve either a nutritional or an osmotic
CC protection function.
CC
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as ics content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```
CC -----
DR EMBL; AB005554; BAA21604.1; -.
DR EMBL; Z99124; CAB16017.1; -.
DR PIR; D70073; D70073.
DR Subtilist; BG11360; csbc.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003863; Sugar_transp.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR Transprot; Transmembrane; Complete proteome.
KW DOMAIN 1
FT TRANSMEM 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 35 1 (POTENTIAL).
FT DOMAIN 36 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 59 2 (POTENTIAL).
FT DOMAIN 60 76 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 77 97 3 (POTENTIAL).
FT DOMAIN 98 104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 105 125 4 (POTENTIAL).
FT DOMAIN 126 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 5 (POTENTIAL).
FT DOMAIN 161 163 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 164 184 6 (POTENTIAL).
FT DOMAIN 185 241 7 (POTENTIAL).
FT TRANSMEM 242 262 8 (POTENTIAL).
FT DOMAIN 263 280 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 281 301 9 (POTENTIAL).
FT DOMAIN 302 308 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 309 329 10 (POTENTIAL).
FT DOMAIN 330 341 11 (POTENTIAL).
FT TRANSMEM 342 362 12 (POTENTIAL).
FT DOMAIN 363 378 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 379 399 13 (POTENTIAL).
FT DOMAIN 400 402 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 403 423 14 (POTENTIAL).
FT DOMAIN 424 461 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 400 401 SA -> RP (IN REF. 1).
SQ SEQUENCE 461 AA; 50140 MW; AD59CB5D6372FE0 CRC64;

Query Match
Best Local Similarity 33.2%; Pred. No. 2.e-34;
Matches 157; Conservative 85; Mismatches 190; Indels 41; Gaps 12;

QY 14 KGNVYASICALLMASVILGYDIGVSGAAMVKKDINIDVQLBILIGLSLYSLF 73
D 2 KKDTRKY--MYFFGALGGLLYGYDTGVISGALLFINNDIPFTLTTEGLVYVSMLLGAIF 59
QY 74 GSFAGARTSDRIGRELTVVFAAVIFFVGSLLMGPAVNYGMLMAGRFVAGVGVGGMIAF 133
D 60 GSALSGTCDRWGRKVVVLSIIFILGALACAFSQTIGMLIASRVLGLVAGVGSTALVP 119
QY 134 VTABISPAASRGFTTTPPEVINTGILLGYLSNFAFARLPPLHGWVLAIGAVPGLL 193
D 120 VYLSEMAPTKIRGTGLTGMNIMIVTGILLAVTVNYLFT--PFE-AWRMNVGLAAVPAVLL 176
QY 194 ALLVFCMPSRWLVKGLRDARAVLEKTSATPEAAERLADKAAAGIPKGLGDGWT 253
D 177 LGIAFMPSRWLVKRGSEEARIMNITH-DPKDIEMELAEKK----- 220
QY 254 VPKQGGGGELOWKKLISLTPPA--VRRILISAVGLHFFQAGSDSVWQYSARLFKSA 311
D 221 -----QGBAEK---KETTGLVKAKWIRPMLLIGVGLAIFQAVGINVIYVFTIFKA 272
QY 312 GITDKNLLGVTCVAGVTKTFILVATFLLDRAGRPLLLISTGGMIVSLICLGLTVA 371
D 273 GLGTSASALG-TMGIGILNVINCITAMILIDRVGRKLLIWGSGVITLSLAALSGVLLT 331
QY 372 GHPDPTKVAWAVALCIASLTSLVIAFFSGICLGPITGVYTSITFIPQLVR--ALGFVAGVASN 429
```

## RESULT 3

```
XYLT LACBR STANDARD; PRT; 457 AA.
ID XYLT LACBR
AC OS2733;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-xylose-proton symporter (D-xylose transporter).
GN XYLT.
OS Lactobacillus brevis.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1580;
RN [1]
SEQUENCE FROM N.A.
RA Chaillou S., Bor Y.-C., Batt C.A., Postma P., Pouwels P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Uptake of D-xylose across the boundary membrane with
CC the concomitant export of a proton (symport system).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF045552; AAC95127.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003863; Sugar_transp.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Symport.
FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 14 34 1 (POTENTIAL).
FT DOMAIN 35 45 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 46 66 2 (POTENTIAL).
FT DOMAIN 67 80 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 101 3 (POTENTIAL).
FT DOMAIN 102 103 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 104 124 4 (POTENTIAL).
FT DOMAIN 125 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 151 5 (POTENTIAL).
FT DOMAIN 152 163 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 164 184 6 (POTENTIAL).
FT DOMAIN 185 243 7 (POTENTIAL).
FT TRANSMEM 244 264 8 (POTENTIAL).
FT DOMAIN 265 280 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 281 301 9 (POTENTIAL).
FT DOMAIN 302 309 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 310 330 10 (POTENTIAL).
FT DOMAIN 331 337 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 338 358 11 (POTENTIAL).
FT DOMAIN 359 379 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 380 400 12 (POTENTIAL).
FT DOMAIN 401 401 EXTRACELLULAR (POTENTIAL).
```

```
FT TRANSMEM 402 422 12 (POTENTIAL).
FT DOMAIN 423 457 CITOPLASMIC (POTENTIAL).
SQ SEQUENCE 457 AA; 49199 MW; E097EB2B67A92F67 CRC64;

Query Match 23.1%; Score 591.5; DB 1; Length 457;
Best Local Similarity 30.6%; Pred. No. 3.4e-32;
Matches 151; Conservative 101; Mismatches 184; Indels 57; Gaps 11;

QY 29 SMASVILGVDMGSAWYIKKQINLTDVQLEIILGILSYLPGSPAGARTSDRIQGR 88
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 14 ALGGLFGYDGTGVSIGAILFIQKQNLGWSQGWVSAVLLGAILGAALIGPSSDRFGR 73
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 89 LTVFAAIVFFVGSLLMGFAVNVGMAGRFVAGVGVGVGMIAFVYTAETISPAASRGL 148
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 74 KULLLSAIFFFVGLGSAFSPFETLIRIILGMVAGAAALIFTYLAELAPSKRTV 133
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 149 TTFPEVINIGILGYLGNFAPARLPHLGNRWMLAIGAVPSGLLALLVFCMPSPRWLV 208
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 134 SSLFQLMVMTGILLAYITNYSF--GFYTGWRWMLGFAAIPAALLFLGGLILPSPRLV 191
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 209 LKGLADARAVLEKTSATPEAAER-LADIKAAAGIPKGLDGVVTVPGKGGGGLQVW 267
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 192 KSGHLEAKRHVLDTNKQDQVANKKEINDIQSA-----KIVSGWSELF 236
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 268 KKLILSPTPAVRRIILLSAVGLHFFQAGSDSVVOYSARLFKSAGITDDNKLGVTC-- 325
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 237 GKM-----VRPSLIIGIGLAIFQQVMGCVLVYVPTFTDVG-----FGVSAALL 282
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 326 ----VGVTKTFILVATFLDRAGRPRLLLISTGGMVLSLCLGSLTVAGHPDVKVW 381
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 283 AHIGIGFIVTAIAVAIMADIRKKNVIGAVGIGSLFVMSIGKMGSGGSGTAALIS 342
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 382 AVALCIASLTVIAFFSIGLGPITGVYTSIEIPLOVRALGFAVGVASNRVTSVISMFTL 441
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 343 VIALIV-----YIAPFSATGPMVMVMEICEVPLNIRGLGNSFASVINWAMVSLTFP 397
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 442 SI-SKAITIGSFLVSGIAAANVFFF-TCLPETRGRLTEENGKFLGMPDTCMAEEADA 500
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 398 SLDDFGGT-GSLFIGYILCFASIFWQKKVPETNRSLD-----IEATLRKGTGDA 450
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 501 AAKEKVELPSSK 513
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 451 A-----ELSTTK 457
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 4
ID ARAE ECOLI STANDARD; PRT; 472 AA.
AC P09830; Q46937;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arabinose-proton symporter (Arabinose transporter).
GN ARAE OR B2841 OR Z4161 OR ECS3698.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
[1]_
RN SEQUENCE FROM N.A.
RX MEDLINE=87115869; PubMed=3543693;
RA Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA Henderson P.J.F.;
RT "Mammalian and bacterial sugar transport proteins are homologous.";
RL Nature 325:641-643 (1987).
[2]
RN SEQUENCE FROM N.A.
RX STRAIN=K12 / JM2433;
RX MEDLINE=88228015; PubMed=2836407;
RA Maiden M.C.J., Jones-Wortimer M.C., Henderson P.J.F.;
RT "The cloning, DNA sequence, and overexpression of the gene araE
RT coding for arabinose-proton symport in Escherichia coli K12.";
```

```
RL J. Biol. Chem. 263:8003-8010 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MGL1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
[4]
RN SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
[5]
RN SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 050952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
[6]
RN PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=84114866; PubMed=6319708;
RA Stoner C., Schleif R.F.;
RT "The araE low affinity L-arabinose transport promoter. Cloning,
RT sequence, transcription start site and DNA binding sites of
RT regulatory proteins.";
RL J. Mol. Biol. 171:369-381 (1983).
CC -!- FUNCTION: Uptake of arabinose across the boundary membrane with
CC the concomitant export of a proton (symport system).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03732; AAA23469.1; -
CC EMBL; X00272; CAA35075.1; -
CC EMBL; U29581; AAB40488.1; -
CC EMBL; AE000368; AAC75880.1; -
CC EMBL; AE005513; AAG57953.1; -
CC EMBL; AP002563; BAB37121.1; -
CC PIR; B26430; B26430.
CC PIR; B91091; B91091.
CC PIR; B85936; B85936.
CC ECoGene; EG10056; araE.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGTRANSPORT.
CC TIGRfams; TIGR00879; SP; 1.
CC PROSITE; PS00850; MFS; 1.
```

DR PROSITE; PS00216; SUGAR TRANSPORT 1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Transpote; Sugar transport; Transmembrane; Inner membrane; Symport;  
KW Complete proteome.  
FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 30 50 1 (POTENTIAL).  
FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 64 84 2 (POTENTIAL).  
FT DOMAIN 85 91 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 92 112 3 (POTENTIAL).  
FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 115 135 4 (POTENTIAL).  
FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 155 175 5 (POTENTIAL).  
FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 179 199 6 (POTENTIAL).  
FT DOMAIN 200 257 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 258 278 7 (POTENTIAL).  
FT DOMAIN 279 297 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 298 318 8 (POTENTIAL).  
FT DOMAIN 319 325 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 326 346 9 (POTENTIAL).  
FT DOMAIN 347 361 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 362 382 10 (POTENTIAL).  
FT DOMAIN 383 404 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 405 425 11 (POTENTIAL).  
FT DOMAIN 426 427 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 428 446 12 (POTENTIAL).  
FT DOMAIN 447 472 CYTOPLASMIC (POTENTIAL).  
FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).  
SQ SEQUENCE 472 AA; 51684 MW; 411990AA41D44393 CRC64;  
Query Match 22.9%; Score 586.5; DB 1; Length 472;  
Best Local Similarity 29.9%; Pred. No. 7.5e-32;  
Matches 147; Conservative 104; Mismatches 197; Indels 43; Gaps 9;  
QY 1 MASDELAKAVPRKGNVKKYASI-CAILASMASVLGVYDIGVMSGAAMWIKKOLNIDVQ 59  
DB 1 WVTINTESALTFRSLURDRRMNMFVSVAAGVLGLDGVIAAGALPFIIDHVLVLSRL 60  
QY 60 LEILIGLSLXLSFGSFAAGTSRDRGLRLVWFAAVIFFVGSLLMGFVAVNYGMLMAGRF 119  
DB 61 QEWVVSMMWGAATGALFNGWLSFRLGRKYSLMAGAILFVLGSGSAFATSVENMLIARV 120  
QY 120 VAGVGVGGMIAPVYTAIEISPAASRGELTTFPEVFNIGILLGYLSNFAFAPLHLG- 178  
DB 121 VLGIAGVGIASVTAPLYSEMASENVGRKMSIMYQLMVTGLGVLAFLSDTAFS----YSGN 176  
QY 179 WRVMLAICAVPSGLLALVFCWPSRWLVKGLADARAVLEKTSATPPEAAERLADIK 238  
DB 177 WRMLGVLLAPVALLIILVFLNPSRWLAKGHIKAEVLMRLTSEKARELNEIR 236  
QY 239 AAAGIPKGLDGVTVPGKEGGGELQWKKLILSPTPAVRILLVAVGLHFFQOASGSD 298  
DB 237 ESKKL-----KQGG-----WALFKINRN--VRRVFLGVLQAMQOFTGMN 275  
QY 299 SVVQYARLFRKSAGITDDNKLGVTCAGVTKTFFILVATFLLDRAGRRPLLISGMI 358  
DB 276 IIMYAPRIFKWAQFTTTEQQMIATLVVGLTFMPTATFVATVVDKGRKPKALKIGFVMA 335  
QY 359 VSLICLG-----SGLTVAGHPDVKVAVAVLACIATSLSYAFFSIGLPIPTGVYTS 411  
DB 336 LGTLVLGCLYQFNGTASSG-----LSW--LSVGMTWMCIAGYAMSAFVVMILCSE 386  
QY 412 IFPLQVRALGFVAGVNASNRVTSVISMFTLSLKAITIGGSFFLYSGIAAVAVVFFTC 471  
DB 387 IQPLKCRDFGTCSTTINWVSNMIGATFLLDLSIGAAGTFWLYTALNIAFVGITFWLI 446  
QY 472 PETGRITLEEM 482  
DB 447 PETKNVILEHI 457

RESULT 5  
ID ITR2 SCHPO STANDARD; PRT; 557 AA.  
AC P87110; P78901;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Myo-inositol transporter 2.  
GN ITR2 OR SPAC20G8.03.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
CX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=968 h90;  
RX MEDLINE=9828285; PubMed=950432;  
RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,  
RA Rusu M., Poitelea M., Edenharter L., Schweingruber M.E.;  
RT "Exogenous inositol and genes responsible for inositol transport are  
RT required for mating and sporulation in Schizosaccharomyces pombe";  
RL Curr. Genet. 33:255-261 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Scourto J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S.E., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Ruster S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Carrutti L., Lowe T., McCombie W.B., Paulsen I., Potashkin J.,  
RA Shipakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe";  
RL Nature 415:871-880 (2002).  
RN [3]  
RP SEQUENCE OF 166-541 FROM N.A.  
RC STRAIN=PR745;  
RX MEDLINE=98162722; PubMed=9501991;  
RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;  
RT "Identification of open reading frames in Schizosaccharomyces pombe  
RT cDNAs";  
RL DNA Res. 4:363-369 (1997).  
CC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to the sugar transporter family.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC

```

DR EMBL; X99105; -, NOT ANNOTATED_CDS.
DR EMBL; Z95334; CAB08597.1; -.
DR EMBL; D89252; BAA13913.1; -.
DR PIR; T38125; T38125.
DR GenDB; Spombe; SPAC208.03; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
DR Transmembrane; Sugar transport.
KW DOMAIN 1 76 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 77 97 1 (POTENTIAL).
FT DOMAIN 98 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 120 2 (POTENTIAL).
FT DOMAIN 121 123 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 124 144 3 (POTENTIAL).
FT DOMAIN 145 157 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 180 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 181 201 5 (POTENTIAL).
FT DOMAIN 202 209 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 210 230 6 (POTENTIAL).
FT DOMAIN 231 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 261 7 (POTENTIAL).
FT DOMAIN 262 367 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 368 388 8 (POTENTIAL).
FT DOMAIN 389 396 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 397 417 9 (POTENTIAL).
FT DOMAIN 418 432 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 433 453 10 (POTENTIAL).
FT DOMAIN 454 468 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 469 489 11 (POTENTIAL).
FT DOMAIN 490 498 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 499 519 12 (POTENTIAL).
FT DOMAIN 520 557 CYTOPLASMIC (POTENTIAL).
FT VARIANT 271 271 N -> H (IN STRAIN PR745).
FT VARIANT 428 428 N -> H (IN STRAIN PR745).
FT VARIANT 436 436 L -> I (IN STRAIN PR745).
FT VARIANT 526 526 S -> F (IN STRAIN PR745).
FT SEQUENCE 557 AA; 61136 MW; 20875EC1B153175 CRC64;

Query Match 22.9%; Score 586; DB 1; Length 557;
Best Local Similarity 31.1%; Pred. No. 9.7e-32;
Matches 154; Conservative 101; Mismatches 202; Indels 38; Gaps 13;

QY 3 SDELAKEPRKGNVKKYASICAILASVASVILGYDVGWGAAMYIKKDLN--ITDVLQ 60
DB 65 ADENEGEFAEKISS--WIVLSAVAGISGLLFGYDTGVISGALAVLSDLGHLSSQOK 122
QY 61 ELILGLTSLYSIFGSGFAGARTSDRIGRLTVVFAVIFVPSGLLGMFAVYVGMAGRFV 120
DB 123 ELITSATSPALISATISGSLADWYGRKRLLCADAIFVIGSVIMAASRNVMVMVGRFI 182
QY 121 AGVGVGWGVIAPVYTAETSPASGGLTTPFEVFINIGILLGYSNFAFARLPHLQWR 180
DB 183 VYGLGLTLIVPMITELAPRLGRLLVIVVITGQLIAYSLNAPEH--VHQWR 240
QY 181 VMLAIGAVPSGLLALVFCMPSPRLVKGSLADARAVLEK--TSATPEFAERLADIK 238
DB 241 IMFGIGAALQGLISLFTSPSPRYLLRHNVKVKILSRHPEAKPAETAYKVSILQ 300
QY 239 AAGIPKGLGDVVTVPGKEQGBELQWKKL--ILSPTPAVRILLSAVGHFFQOASGS 297
DB 301 -----EGVKVDFP-----EGNKQFHFFHSUKVLFTVPSNRSLSFGCFQWFOQSGT 348
QY 298 DSVVOYSARLFKASGITDDNKLGLGVTCAGVGVTKTFEILVATFLDRAGRRLFLISTGCM 357

```

## RESULT 6

```

ARAE KLEOX
ID -ARAE KLEOX STANDARD; PRT; 472 AA.
AC P45598;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arabinose-proton symporter (Arabinose transporter).
GN ARAE.
OS Klebsiella oxytoca.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=571;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8017;
RA MEDLINE=95394866; PubMed=7665532;
RX Shatwell K.P., Charalambous B.M., McDonald T.P., Henderson P.J.F.;
RT "Cloning, sequencing, and expression of the araE gene of Klebsiella
RT oxytoca 8017, which encodes arabinose-H+ symport activity.";
RL J. Bacteriol. 177:5379-5380(1995).
CC -!- FUNCTION: Uptake of arabinose across the boundary membrane with
CC the concomitant export of a proton (symport system).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X79598; CA456110.1; -.
CC PIR; S47089; S47089.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transpt.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGRTRNSPORT.
CC TIGRFAMs; TIGR00879; SP; 1.
CC PROSITE; PS50850; MFS; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 30 50 1 (POTENTIAL).
FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 64 84 2 (POTENTIAL).
FT DOMAIN 85 91 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 92 112 3 (POTENTIAL).
FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 115 135 4 (POTENTIAL).
FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 5 (POTENTIAL).
FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).

```



FT	TRANSMEM	179	199	6 (POTENTIAL).	RN
FT	DOMAIN	200	257	CYTOPLASMIC (POTENTIAL).	RP
FT	TRANSMEM	258	278	7 (POTENTIAL).	RC
FT	DOMAIN	279	297	PERIPLASMIC (POTENTIAL).	RX
FT	TRANSMEM	298	318	8 (POTENTIAL).	RA
FT	DOMAIN	319	325	CYTOPLASMIC (POTENTIAL).	RA
FT	TRANSMEM	326	346	9 (POTENTIAL).	RA
FT	DOMAIN	347	361	PERIPLASMIC (POTENTIAL).	RA
FT	TRANSMEM	362	382	10 (POTENTIAL).	RA
FT	DOMAIN	383	404	CYTOPLASMIC (POTENTIAL).	RA
FT	TRANSMEM	405	425	11 (POTENTIAL).	RA
FT	DOMAIN	426	446	PERIPLASMIC (POTENTIAL).	RA
FT	TRANSMEM	428	446	12 (POTENTIAL).	RA
FT	DOMAIN	447	472	CYTOPLASMIC (POTENTIAL).	RA
SQ	SEQUENCE	472 AA;	51732 MW;	410021E1BEE3D96E CRC64;	
Query Match					
Best Local Similarity					
Matches 142; Conservative 96; Mismatches 186; Indels 42; Gaps 8;					
Qy	25	AIALMASGVLGIDGVNMGAAVYIKKDLNLTVDQLEILIGLSLFLGSGFAGARTSDR	84		
Db	26	SIAAVALGLGLDGLVAGALPFTIDHVLSSRLQEWVSSMMGLGAIGALFNGLSFR	85		
Qy	85	IGRLTVFAVIFVPGSLMGFAVNYGMAGRFVAGVGVGGMIAPVYTAISPAAS	144		
Db	86	LGRKYSLMVGAVLFVAGSVGSFATSVEMLLVARIVLGVAVGIASYPAPLYLSEMAENV	145		
Qy	145	RGELTFPEVINIGILGLYLSNFAFALPHLG-WRVMALGAVPSGLLALLVFCPES	203		
Db	146	RKMISMVQLMVTIGVNAFLSDTAFS-----ISGNWRMLGVLPALPAVLLILVIFLPS	201		
Qy	204	PRWLVLKGLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGDVVTPVPGKEQGGGE	263		
Db	202	PRWLAEGKRVHAEVLMRLDTSEKARDELNEIRELSKL-----KQGG--	245		
Qy	264	LOVWKULSTPAVRRLILLSAVGLHFFQOASGSDSVVVSARLFKSGAGITDDNKLGV	323		
Db	246	---W--ALFKYNNVRRAVFLGMLLQAWQOFTGNWIIIMYAPRIFKVGAGTTTBEQQMVAT	300		
Qy	324	CAVGVTKTFLLVATFLDLAGRPLLLISITGGMIVSLICLG-----SGLTVAGHPD	376		
Db	301	LVVGLTFMEAFIAVFTVDKAGRPALKIGSVNAIGTLVLGYCLMQFDNGTASSG----	356		
Qy	377	TKVAVAVALCTASLSYIAFTSISGLGPIGTGYTTEIPLQVRALGFVAVASNRVTSVAVI	436		
Db	357	--LSW--LSVGMTCMCIAGVAMSAAPVVMVILCSEIQPLKCRDFGTCSTTTNWNVMII	411		
Qy	437	SWTFLSKATITIGSFPLYSGIAAVAVWRFFTCCLPBTGRGRTLEEM	482		
Db	412	GATFLLDALGAAGITFWLYTALNVAFIGVTFWLPETKNTLLEHI	457		
RESULT 7					
ID	YFIG_BACSU	STANDARD;	PRT;	482 AA.	
AC	P54723;				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Hypothetical metabolite transport protein yfig.				
GN	YFIG OR BSU08260.				
OS	Bacillus subtilis.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxID=1423;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=168;				
RX	MEDLINE=96262713; PubMed=8704981;				
RA	Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;				
RT	"Determination of a 12 kb nucleotide sequence around the 76 degrees				
RT	region of the Bacillus subtilis chromosome."				
RL	Microbiology 142:1417-1421(1996).				

[2]

SEQUENCE FROM N.A.

STRAIN=168;

MEDLINE=98044033; PubMed=9384377;

Kunest F., Ogasawara N., Moser I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Boursier L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choe S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Errington J., Fabret C., Ferrari E., Foulger D., Fritzsche K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerz-Blanchard M., Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetalle D., Porwollik S., Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Roche A., Roche B., Rose M., Sadaie Y., Sato T., Scanlon A., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.; RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*."

RL Nature 390:249-256(1997).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -1- SIMILARITY: Belongs to the sugar transporter family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; D50543; BAA09111.1; -

DR EMBL; Z99108; CAB12655.1; -

DR PIR; B69803; B69803.

DR Subtilist; BG11854; yfig.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub\_transporter.

DR InterPro; IPR005829; Sug\_transporter.

DR InterPro; IPR003663; Sugar\_transp.

DR Pfam; PF00083; sugar\_tr\_1

DR PRINTS; PR00171; SUGRTRANSPORT.

DR TIGRfams; TIGR00879; SP; 1.

DR PROSITE; PS00850; MFS; 1.

DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; FALSE\_NEG.

DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.

KW Hypothetical protein; Transport; Transmembrane; Complete proteome.

FT DOMAIN 1 29

FT CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 30 50

FT DOMAIN 51 59

FT EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 60 80

FT DOMAIN 81 92

FT CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 93 113

FT DOMAIN 114 120

FT EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 121 141

FT DOMAIN 142 155

FT CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 156 176

FT DOMAIN 177 184

FT EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 185 205

FT CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 206 220

FT EXTRACELLULAR (POTENTIAL).

FT DOMAIN 206 263 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 284 284 7 (POTENTIAL).  
 FT DOMAIN 285 301 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 302 322 8 (POTENTIAL).  
 FT DOMAIN 323 331 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 332 352 9 (POTENTIAL).  
 FT TRANSMEM 333 373 10 (POTENTIAL).  
 FT DOMAIN 374 400 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 401 421 11 (POTENTIAL).  
 FT DOMAIN 422 423 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 424 444 12 (POTENTIAL).  
 FT DOMAIN 445 482 CYTOPLASMIC (POTENTIAL).  
 SQ SEQUENCE 482 AA; 52756 MW; 9085104F4C48EE01 CRC64;

Query Match 22.2%; Score 568.5; DB 1; Length 482;  
 Best Local Similarity 31.1%; Pred. No. 1.2e-30;  
 Matches 146; Conservative 93; Mismatches 191; Indels 39; Gaps 10;

QY 26 ILASNASVILGVDIGVMSGAAMVYKK--DLNITDVLQLEILIGILSLYSLFGSFAGARTSD 83  
 DB 26 LVSTGGLFGYDTGVINGALPFPATAGLNLTPTVEGLVASSLLGAFGAMFGRLSD 85  
 QY 84 RIGRLTVFAAVIFVGSLLMGPAVNYCMAGRFVAGVGVGGMTAPVYTAISPA 143  
 DB 86 RHGRKRTILYALLFAATLGTCTSPNASVMIAFRFLGLAVGCASVTVPVTFLEISPAE 145  
 QY 144 SRGFLTPPEVINIGLLYLSNFAFARLPPLHLG-----WRVMLAIGAVPSGLLALIVF 198  
 DB 146 RGRIVTQNELMIVLQLLAYTFN---AIGSTMGESANWRVLMVIALTPAVLVFQML 202  
 QY 199 CMPESPRLVLKRLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGDVVTVPQKE 258  
 DB 203 IVPESPRMLAAGRGDALLRVLQI-REDSQAQKEIKIA-----IEG-----TAKK 250  
 QY 259 QGGGLQWKKLILSFTPAVRILLSAVGLHFFQASGSDSVQVSARLFKSAGITDKNK 318  
 DB 251 AGFHDQ-----EPWIRILFGIGIAVQOITGNSIMYGTILREAGFQTEAA 301  
 QY 319 LIGVTCAGVTKTFILNATFLDRAGRRPLLLISTGGMVSLICLGSGLTVAGHHPDTK 378  
 DB 302 LIG-NIANGVSVIAVIGILWLGKVRRLPMLIIQIGTWTALLIGILSVLECTP--- 357  
 QY 379 VMAVALCTAS-TLSYIAFFSLGDPITGVYTSIFPLQVRLGFAVGAVNRVTSAVIS 437  
 DB 358 ---ALPYVVLSTILFLAQQTATSTVTLMLSEIFPMHVRLGLGMGISTFCLWTANFLIG 414  
 QY 438 MTFLSLSKAITGGFFLYSGIAAVAVVFFFCPLPETRGRTELEMGKLF 486  
 DB 415 FIFPILLNHIGSATFFIFVAVNAILALFVKYVETKGRSLQLEHSP 463

## RESULT 8

ID IT1 YEAST STANDARD; PRT; 584 AA.  
 AC P30805;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Myo-inositol transporter 1.  
 GN IT1 OR YDR497C OR D9119.3.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91250431; PubMed=2040626;  
 RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;  
 RT "Isolation and characterization of two distinct myo-inositol  
 transporter genes of Saccharomyces cerevisiae."  
 RL J. Biol. Chem. 266:11184-11191 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,  
 BA Bero A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,  
 RA Hunkle-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,  
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,  
 RA Minant A., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAJOR TRANSPORTER FOR MYO-INOSITOL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: Belongs to the sugar transporter family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL; D90352; BAA14366.1; -;  
 EMBL; U33057; AAB64939.1; -;  
 PIR; S69555; S69555.  
 GERMOnline; 140989;  
 SGD; S0002905; ITR1.  
 DR GO; GO:0005365; P:myo-inositol transporter activity; IMP.  
 DR GO; GO:0015798; P:myo-inositol transport; IMP.  
 DR InterPro; IPR007114; MFS.  
 DR InterPro; IPR005828; Sub transporter.  
 DR InterPro; IPR005829; Sug transporter.  
 DR InterPro; IPR003663; Sugar\_transpt.  
 DR Pfam; PF00083; sugar\_crr; 1.  
 DR PRINTS; PR00171; SUGTRNSPORT.  
 DR TIGRFAMS; TIGR00879; SP; 1.  
 DR PROSITE; PS00850; MFS; 1.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 2.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 DR Transmembrane; Sugar transport; Glycoprotein.  
 KW DOMAIN 1 81 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 82 102 1 (POTENTIAL).  
 FT DOMAIN 103 129 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 130 150 2 (POTENTIAL).  
 FT DOMAIN 151 163 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 164 184 3 (POTENTIAL).  
 FT DOMAIN 185 186 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 187 207 4 (POTENTIAL).  
 FT DOMAIN 208 215 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 216 236 5 (POTENTIAL).  
 FT DOMAIN 237 246 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 247 267 6 (POTENTIAL).  
 FT DOMAIN 268 349 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 350 370 7 (POTENTIAL).  
 FT DOMAIN 371 376 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 377 397 8 (POTENTIAL).  
 FT DOMAIN 398 400 9 (POTENTIAL).  
 FT TRANSMEM 401 421 10 (POTENTIAL).  
 FT DOMAIN 422 441 11 (POTENTIAL).  
 FT TRANSMEM 442 462 12 (POTENTIAL).  
 FT DOMAIN 463 485 11 (POTENTIAL).  
 FT TRANSMEM 487 507 12 (POTENTIAL).  
 FT DOMAIN 508 510 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 511 531 13 (POTENTIAL).  
 FT DOMAIN 532 584 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 43 44 TL -> HI (IN REF. 1).  
 SQ SEQUENCE 584 AA; 63569 MW; 4254330A102DC65 CRC64;

Query Match 22.1%; Score 565; DB 1; Length 584;  
 Best Local Similarity 29.5%; Pred. No. 2.5e-30;  
 Matches 150; Conservative 109; Mismatches 183; Indels 66; Gaps 13;

QY 27 LASMASVILGVDIGVMSGAAMVYKKDLN--ITDVLQLEILIGILSLYSLFGSFAGARTSD 83  
 DB 92 VASISGMFGYDTGYTGISSALISIGTLDHKLVTYGEKEIVTAATSLGALITSIFACTAAD 151

```
QY 84 RIGRLTVVFAAIVFVGLMGMFAVNYGMLMAGRFVAGVGVGYGMIAPVVTAISPA 143
Db 152 IFKCKCLGNSLMFVIGAILQVSAFTFQMAVGRLLMGVGVGSLIAPLFIETIAPK 211
QY 144 SRGLFTTFEVFINIGIL-----LGYLSNFAFAPARLPLHLGWRVMLAIGAVPSGLL 195
Db 212 IRGLTVINSMLTGGQVAYGCGAGLNIVN-----GWRILVGLSIPTAVQPT 261
QY 196 LVFCMPESRWLVKRLDARAVLEKT-SATPEEAERLADIKAAAGIPKGLDGDVVT- 253
Db 262 CLCFLPDTTRYVMKGLDARATEVLKRSYTDTSSEIIERKVE-----ELVTL 308
QY 254 ---VPGKEQGGELQWKKL-ILSPTPAVRILLISAVGLHFFQOAGSDSVVQYSARLF 309
Db 309 NOSIPQNV---PEKVNVIKELHTVPSNRLIITCGQLAQOQFTGWSNLMFYSGTIFE 365
QY 310 SAGITDDNKLGVTCVAVGVTKFFILVATFLDRAGRRPLLIISGGMIVSLICLGSQT 369
Db 366 TVGFKNS---AVSIIVSGTNEIFTLVAFSSIDKIGRRITILLIGLPGMTALVV---CS 418
QY 370 VAGHPDTPKVAVALCIASITLS-----YIAPFSGLGPIGTGVYTSIEFPLOV 417
Db 419 IAFHLGKFDGAVAVVSSGSSFGWIVIVFIIVFAFALGIGTVPW-QQSELPQNV 477
QY 418 RALGFVAVGASNRVTSVISMVFLSLSKAITIGGSFELYSGIAAVAWVFFFCLPETRGR 477
Db 478 RIGTSYATATNAGSLVIASTFLMLQNLITAGTFAPFAGLSCLSTIFCYFELSG 537
QY 478 TLEEMKLGPMPTGMABEAEADAKEK 505
Db 538 ELEEVQTL---KGFNFKASKALAKR 562

RESULT 9
ITR2 YEAST
ID ITR2 YEAST STANDARD; PRT; 612 AA.
AC P30606;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 2.
GN ITR2 OR YOL103W OR HRB612.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91250431; PubMed=2040626;
RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT "Isolation and characterization of two distinct myo-inositol
RL transporter genes of Saccharomyces cerevisiae.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96076631; PubMed=7502582;
RA Vandenbol M., Durand P., Portetelle D., Hilger F.;
RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV
RT including the Ty1-H3 retrotransposon, the sufl(+) frameshift
RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a
RT delta element.";
RL Yeast 11:1069-1075 (1995).
CC -!- FUNCTION: MINOR TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
```

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```
CC
DR EMBL; D90353; BAA14367.1; -.
DR EMBL; Z48149; CAA88159.1; -.
DR EMBL; Z74845; CAA99119.1; -.
DR PIR; B40538; B40538.
DR Germonline; 143525; -.
DR SGD; S0005463; ITR2.
DR GO; GO:0005365; F:myo-inositol transporter activity; IMP.
DR GO; GO:0015738; P:myo-inositol transport; IMP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 107 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 108 128 1 (POTENTIAL).
FT DOMAIN 129 155 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 156 176 2 (POTENTIAL).
FT DOMAIN 177 182 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 183 203 3 (POTENTIAL).
FT DOMAIN 204 212 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 213 233 4 (POTENTIAL).
FT DOMAIN 234 241 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 242 262 5 (POTENTIAL).
FT DOMAIN 263 272 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 273 293 6 (POTENTIAL).
FT DOMAIN 294 362 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 363 383 7 (POTENTIAL).
FT DOMAIN 384 402 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 403 423 8 (POTENTIAL).
FT DOMAIN 424 426 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 427 447 9 (POTENTIAL).
FT DOMAIN 448 467 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 468 488 10 (POTENTIAL).
FT DOMAIN 489 512 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 513 533 11 (POTENTIAL).
FT DOMAIN 534 536 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 537 557 12 (POTENTIAL).
FT DOMAIN 558 612 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 612 AA; 67041 MW; 48E733102BAD7298 CRC64;
```

Query Match 22.0%; Score 563; DB 1; Length 612;  
Best Local Similarity 29.3%; Pred. No. 3.6e-30;  
Matches 149; Conservative 108; Mismatches 184; Indels 68; Gaps 16;

```
QY 27 LASMASVILGYDIGVMSGAMVYKKDLN---ITDVOLEILIGLSLYSLFGSPAGARTSD 83
Db 118 VASISGMFGYDGTGYISSALISINRDLNKNLTYGEKELITATSLGALITSVGAGTAAD 177
QY 84 RIGRLTVVFAAIVFVGLMGMFAVNYGMLMAGRFVAGVGVGYGMIAPVVTAISPA 143
Db 178 VFGRRCPLMFSNLMFLIGAILQITAKHPQWMAAGRLMGFGVGIGLSLFIETIAPK 237
QY 144 SRGLFTTFEVFINIGILGY-----LSNFAFARLPLHLGWRVMLAIGAVPSGLL 195
Db 238 IRGLTVINSMLTGGQLIAYGCGAGLNHVKN-----GWRILVGLSIPT-VLQF 286
QY 196 LVFC-MPESPRWLVKRLDARAVLEKTSATPE-----EAAERLADIKAAAGIPKGLD 249
Db 287 SFFCFUFDTPRYVMKGLDKRAVNLKRSYNTDEIDQKVEELSSLNQ----- 336
QY 250 DVVTFPGKEGGELQWKKL-ILSPTPAVRILLISAVGLHFFQOAGSDSVVQYSARLF 308
Db 337 ---SIFGKNP---ITKFWNVKXELHTVPSNFRALLIGCLQAIQOFTGWSNLMFYSGTIF 390
```

```

QY 309 KSAGITDKNLLGVTCVAVGVTFFILVATFLLDRGRPLLLISTGGMIYSL-ICL--- 364
DQ 391 ETGVGFNS--AVSIIISVGNFVFTLFAFFCIDKIGRRIYLLIGLPGMTVALVICAIF 447
QY 365 -----GSGUTVAGHPDVKVAVAVLCTASTLSYIAFFSGLGPIGTGYTSEIPLQ 416
DQ 448 HFLGLKFGADAVAS---DGSSNGIVI-IVFIIVYAFVAGLGITGVEM-QQSELPPON 502
QY 417 VREALFVAVGNRVTSVISMTFLSKATIGSFPLYSGIAVAVVFTCLPETR 476
DQ 503 VGVGVTSYATATNAGSLVIASTFLTMQNIHTGTSTFFAGVACLSTFYCFYCPBUSG 562
QY 477 RTLEMGKLFMGPDTCMAEADAAKEX 505
DQ 563 LELEEVQTL---KQGNFKASKALAKR 588

RESULT 10
ID -GALP-ECOLI STANDARD; PRT; 464 AA.
AC P37021; 1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Galactose-proton symporter (Galactose transporter).
GN GALP OR B2943 OR C3529.
OS Escherichia coli, and
OS Escherichia coli 06.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562, 217992;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA Roberts P.E.;
RL Thesis (1992), University of Cambridge, U.K.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O4:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.I.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
CC -1- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: Belongs to the sugar transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DQ EMBL; U28377; AAA69110.1; --
DQ EMBL; AE000377; AAC75980.1; --
DQ EMBL; AE016766; AA881977.1; ALT_INIT.
DQ PIR; F65079; F65079.

```

```

EcoGene; EG12148; galP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar_cr; 1.
DR PRINTS; PRO0171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 1 (POTENTIAL).
FT DOMAIN 37 56 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 84 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 85 105 3 (POTENTIAL).
FT DOMAIN 106 112 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 113 133 4 (POTENTIAL).
FT DOMAIN 134 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 5 (POTENTIAL).
FT DOMAIN 161 171 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 172 192 6 (POTENTIAL).
FT DOMAIN 193 250 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 251 271 7 (POTENTIAL).
FT DOMAIN 272 290 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 291 311 8 (POTENTIAL).
FT DOMAIN 312 321 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 322 342 9 (POTENTIAL).
FT DOMAIN 343 351 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 352 372 10 (POTENTIAL).
FT DOMAIN 373 394 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 395 415 11 (POTENTIAL).
FT DOMAIN 416 437 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 438 464 12 (POTENTIAL).
FT SEQUENCE 464 AA; 50982 MW; 07E08935BD8E3F8E CRC64;
Query Match 21.4%; Score 547; DB 1; Length 464;
Best Local Similarity 28.8%; Pred. No. 3.1e-29;
Matches 141; Conservative 96; Mismatches 205; Indels 48; Gaps 9;
QY 11 EPRKGNVKYAS---ICAILASMASVILGYDVGMSGAAMYIKDLNITDVQLEILIGIL 67
DQ 3 DAKQGRSNKAMTFVC-FLAALAGLFGLDIGVIAGALPFIADFPQITSHTEWVSSM 61
QY 68 SLYSLFGSFAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMAGRFVAGVGVY 127
DQ 62 MFGAAVAGVGWLSFKLGRKSLMIGAILFVAGSLFSAAPXVVEVLLSRVLLGAVGV 121
QY 128 GCMAPYVTARISPAASRGILTPPEVFINIGLLVLSNFARLPLHLG-WRVMLAIG 186
DQ 122 ASYTAPLYLSIAPEKIRGSMISWYQMILIGILCAVLSDTAFS----YTGAWMLGLVI 177
QY 187 AVPSGLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEAERLADIKAAAGPKG 246
DQ 178 IIPALLLIGVFFLPDSFRFAKRRFVDAERVLLRLRDTSAKAKRELDIRESLQV--- 234
QY 247 LDGDVTVPGKEQGGELQWKKLLLSPTPAVRILLSAVGLHFFQOASGSDSVQYSAR 306
DQ 235 -----KQSG-----W--ALFKENSNFRAVFLGULLQWQOFTGMVIMYAPK 276
QY 307 LFKSAGITDDNKLGLGTCVAVGVTFFILVATFLLDRGRPLLLISTGGMIYSLICLS 366
DQ 277 IFELAGYNTTEQMGWTVIGLTVNLFATFIAGLVDWRGRKPTLTGLFLVMAAGMVLGT 336
QY 367 GLTVAGHPDVKVAVAVLCTASTLSYIAFFSGLGPIGTGYTSEIPLQVRLGFAVG 426
DQ 337 MMHIGHSFS-----AQYFAIAMLMTFVGFAMSGAGLFWLVCSEIQPKGRDFGTCST 391
QY 427 ASNRVTSAVISMTFLSLSKAITIGSFPLYSGIAVAVVFTCLPET----- 474

```

```
Db 392 ATNWIAMIVGATFLMNTLGNANTFWYAAALNVILLTLLWLPETKHVSLEHIERNL 451
Qy 475 -RGRTLEENG 483
Db 452 MXGRKLREIG 461

RESULT 11
ID ITR1_SCHPO
AC Q10286; STANDARD; PRT; 575 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 1.
GN ITR1 OR SPAC4F8.15 OR SPAC7D4.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RX SEQUENCE FROM N.A.
RC STRAIN=968 h90;
RX MEDLINE=98228265; PubMed=9560432;
RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
RA Rusu M., Poitelea M., Edenharter L., Schweingruber M.B.;
RT "Exogenous inositol and genes responsible for inositol transport are
RL required for mating and sporulation in Schizosaccharomycetes pombe.";
RN Curr. Genet. 33:255-261(1998).
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leathers S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Woestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutini L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
```

```
DR EMBL; X98622; CAA67211.1; -.
DR EMBL; Z98530; CAB11061.1; -.
DR EMBL; Z99532; CAB16718.1; -.
DR PIR; T43400; T43400.
DR GeneDB SPombe; SPAC4F8.15; -.
DR GO; GO:000747; P:conjugation with cellular fusion; ISS.
DR GO; GO:0006629; P:lipid metabolism; ISS.
DR GO; GO:0007165; P:signal transduction; ISS.
DR GO; GO:0007151; P:sporulation (sensu Saccharomycetes); ISS.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 86
FT TRANSMEM 87 107
FT DOMAIN 108 129
FT TRANSMEM 130 150
FT DOMAIN 151 156
FT TRANSMEM 157 177
FT DOMAIN 178 186
FT TRANSMEM 187 207
FT DOMAIN 208 215
FT TRANSMEM 216 236
FT DOMAIN 237 246
FT TRANSMEM 247 267
FT DOMAIN 268 349
FT TRANSMEM 350 370
FT DOMAIN 371 376
FT TRANSMEM 377 397
FT DOMAIN 398 400
FT TRANSMEM 401 421
FT DOMAIN 422 441
FT TRANSMEM 442 462
FT DOMAIN 463 486
FT TRANSMEM 487 507
FT DOMAIN 508 510
FT TRANSMEM 511 531
FT DOMAIN 532 575
FT CARBOHYD 432 432
SQ SEQUENCE 575 AA; 62757 MW; 3B7C5EFP86C596AE CRC64;

Query Match 21.3%; Score 546; DB 1; Length 575;
Best Local Similarity 30.8%; Pred. No. 4.5e-29;
Matches 153; Conservative 86; Mismatches 206; Indels 52; Gaps 12;

Qy 19 KYASICALLASMASVILGYDIGVMGSAAMVKKDL---NITDVQLBILIGILSLYSLFGS 75
Db 85 KVMVLAFAAGIGGLFGYDTGVISGALVVIGTSLGGHELTNGGKEFITSATSLGALLGG 144
Qy 76 FAGARTSDRIGRLTVFVAFAVFFVCSLLMGFAVNYGMLMAGRFVAGVGVGGMTAPVY 135
Db 145 ILAGALADPFGRKPVIAIASIIIVGSIQVTAHLHWMIVGSRFVIGWGVIASLIPLY 204
Qy 136 TAEISPAARGFLTTPEPEFINIGILGLVSNFAPARLPHLGWRVLAIGAIVPSGLLAL 195
Db 205 LSLIAPSKIRGLVILVILLITAGVIAICIDTAPEH--VHNGWRWVGLWAPAFQFLF 262
Qy 196 LVCPMPESPRWLVKRLADARAVLEK--TSATPPEAERLADIKAAAGIPKGLDGVVT 253
Db 263 ILIWPESPRLLVKKERSQEAYNLTARIYPTAFPEIKTKLYLQ--EGVRDPFSG---- 316
Qy 254 VPKEGGGBGLQWKKLI-----LSPTPAVRRLISAVGLHFFQAGSGSDSVVQVYARLF 308
Db 317 -----SRWQKIVTKFELYPNPNFALILACGLQAMQQLSGNSLWYFSTTIF 365
Qy 309 KSAGITDDNKLGLVTCVAGVTKTFFILVATFLLDRAGRPLLLISTGGMIVSLICLSGL 368
```

Db 366 EVVGF---NNPTATGLIIAATNFVETIVAFGVDFGRILLVWGMIAALIV---C 418  
 QY 369 TVAGHH-----PDKWAVALCIASLSVIAFFSGLGPIITGVYTSFIPLOVR 418  
 Db 419 AVAFHFLPKDENGNTSGOSNWAIVLI-SMIVTVASVAGLGNLPW-QQSEIFPMSVR 476  
 QY 419 ALGFVAVGASRVNTSAVISMTLSKAITGGSFLLYSGIAAVAWVFFTCLPETRGR 478  
 Db 477 GLGTGMSAVWAGNIGASFLTMSLTPGTGTALYGLICFLGWLGFALCYPLDTDT 536  
 QY 479 LEMKGL-----FGMPDT 491  
 Db 537 IBEIGELLXHGFGVRES 553

RESULT 12  
 ID\_XYLE\_ECOLI STANDARD; PRT; 491 AA.  
 AC P09098;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE D-xylose-proton symporter (D-xylose transporter).  
 GN XYLE OR B4031 OR 25639 OR ECS5014.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=88007632; PubMed=2820984;  
 RA Davis E.O., Henderson P.J.F.;  
 RT "The cloning and DNA sequence of the gene xyle for xylose-proton  
 symport in Escherichia coli K12."  
 RL J. Biol. Chem. 262:13928-13932(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87115869; PubMed=3543693;  
 RA Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,  
 RT Henderson P.J.F.;  
 RL "Mammalian and bacterial sugar transport proteins are homologous."  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MGI1655;  
 RX MEDLINE=94089392; PubMed=8265357;  
 RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,  
 RA Daniels D.L.;  
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the  
 region from 89.2 to 92.8 minutes."  
 RL Nucleic Acids Res. 21:5408-5417(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grothbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."  
 RL Nature 409:529-533(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,

Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12."  
 RL DNA Res. 8:11-22(2001).  
 RN [6]  
 RP SEQUENCE OF 1-192 FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=88234001; PubMed=2836810;  
 RA Francoz E., Dassa E.;  
 RT "3' end of the malBFG operon in E. coli: localization of the  
 transcription termination site."  
 RL Nucleic Acids Res. 16:4097-4109(1988).  
 CC -1- FUNCTION: Uptake of D-xylose across the boundary membrane with the  
 concomitant export of a proton (symport system).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -1- INDUCTION: By xylose.  
 CC -1- MISCELLANEOUS: E. COLI HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT  
 ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM  
 WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS  
 INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYLE SYSTEM THAT USES A  
 HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC  
 SHOCK.  
 CC -1- SIMILARITY: Belongs to the sugar transporter family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; J02812; AAA79016.1; --  
 CC EMBL; U00006; AAC43125.1; --  
 CC EMBL; AE000476; AAC77001.1; --  
 CC EMBL; AE005636; AAC59230.1; --  
 CC EMBL; AP002568; BAB38437.1; --  
 CC EMBL; X06663; CAA23863.1; --  
 CC PIR; A26430; A26430.  
 CC PIR; B86096; B86096.  
 CC PIR; F91255; F91255.  
 CC EcoGene; EG1076; xyle  
 CC InterPro; IPR007114; MFS.  
 CC InterPro; IPR005828; Sub-transporter.  
 CC InterPro; IPR005829; Sugar transporter.  
 CC InterPro; IPR003663; Sugar transprt.  
 CC Pfam; PF00083; sugar tr; 1.  
 CC PRINTS; PR00171; SUGRTNSPORT.  
 CC TIGRfams; TIGR00879; SP; 1.  
 CC PROSITE; PS00850; MFS; 1.  
 CC PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 CC PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;  
 KW Complete proteome.  
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 10 30 1 (POTENTIAL).  
 FT DOMAIN 31 55 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 56 76 2 (POTENTIAL).  
 FT DOMAIN 77 89 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 90 110 3 (POTENTIAL).  
 FT DOMAIN 111 133 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 134 154 4 (POTENTIAL).  
 FT DOMAIN 155 165 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 166 186 5 (POTENTIAL).  
 FT DOMAIN 187 200 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 201 221 6 (POTENTIAL).  
 FT DOMAIN 222 272 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 273 293 7 (POTENTIAL).  
 FT DOMAIN 294 312 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 313 333 8 (POTENTIAL).  
 FT DOMAIN 334 343 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 344 364 9 (POTENTIAL).  
 FT DOMAIN 365 PERIPLASMIC (POTENTIAL).

```
FT TRANSMEM 370 390 10 (POTENTIAL).
FT DOMAIN 391 407 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 408 428 11 (POTENTIAL).
FT DOMAIN 429 442 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 443 463 12 (POTENTIAL).
FT DOMAIN 464 491 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 64 64 A -> V (IN REF. 6).
SQ SEQUENCE 491 AA; 53608 MW; 2AF1AF9756C0B722 CRC64;

Query Match 21.3%; Score 544; DB 1; Length 491;
Best Local Similarity 29.9%; Pred. No. 5.2e-29;
Matches 158; Conservative 93; Mismatches 183; Indels 94; Gaps 18;

QY 17 NVKVASICAILASMASVILGIDGVMGAAMVYK-----KDLNIDTDVQLLEIILGILSYSL 72
DB 6 NSSYFSTLIVATLGLGLFGDVTAVISGTVESLNTVFVAPQNLSESAANSLGFCVASAL 65
QY 73 FGSFAGAR-----TSDRIGRRVTVFAAIVFVGLS-----LMGF-AVN-----YGLMAG- 117
DB 66 IGCITGGALGGYCSNRFGRDRSLKAAVLFTISGVSAWPELGFTSINPDNTPVYLAGY 125
QY 118 -----RFVAGVGVGGMIAVPVYTAISPAASEGFLTTPEVFINIGILGILSYLNEAF 170
DB 126 VPEFVIYRIIGIGVGLASMLSPYIAELAPAHIRKLVSPNQFAITPQQLLVYCVNYFI 185
QY 171 ARLP-----LHL-GWRVMLAIGAVPSGLLALLVFCMPESPRWLVLKRLADARAVLEK--- 222
DB 186 ARSGDASWLNTDGRWYMFASECIPALLFLMLLYTVPSPRWLMRGKQEQAEGLRKIMG 245
QY 223 -TSATPEAAERLADIKAAIGPKLGDVDVTPKQEGGGLQVKKLILSPFPAVRI 281
DB 246 NTLAP-----QAVQIKHSLD-----HGRKTGG-----RL 270
QY 282 LLSAVG-----LHFFQAGSDSVQVYARLPSAGITDDNKLGLVTCVAVGVTKTFI 334
DB 271 LMEFGVIVIGVMSLIFQGFVGINVLYYAEVFKTLGASDIAL-CTIIVGVINLTFT 329
QY 335 LVATPELLDRAGRPLLLISTGMIVSLICLGSGLTVAGHHPTKVANAVALCIASITSYI 394
DB 330 VLAIMTVDFKFKPLQIICALGMAIGMPSLGTAFY-----TOAPGIVAL--LSMLFYV 380
QY 395 AFFSIGLGPITGVYSEIFPPLQVRLGFAVGAVSNRVTSAVISMTELSLKAITI----- 449
DB 381 AAFAPSWGVCWLLSEIFPNAIRKALAIYAVQWLANFYVSWTFPMDDKSNLVAHFH 440
QY 450 -GGSFYXGIAAVAWVFFTCLPETRGRTLEEMGKLFMPDGTGMAEE 496
DB 441 NGFSWYIYCGVLAALFMKFWFVPTKTKTLEELW-EPETKTKTQ 487

RESULT 13
ID ARAE BACSU STANDARD; PRT; 464 AA.
AC P96710;
DC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Arabinose-proton symporter (Arabinose transporter).
GN ARAE OR BSU33960.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]_TaxID=1423;
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
Guiseppi G., Guy B.J., Haga K., Haeich J., Harwood C.R., Henaut A.,
Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
Srokin A., Taconelli E., Takagi T., Takahashi H., Takemaru K.,
Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
"the complete genome sequence of the Gram-positive bacterium Bacillus
subtilis";
RL Nature 390:249-256 (1997).
RN [2]
RP SEQUENCE OF 1-223 FROM N.A.
RC STRAIN=168;
RX MEDLINE=97197523; PubMed=9045819;
RA Sa-Nogueira S., Mota L.J.;
RT "Negative regulation of L-arabinose metabolism in Bacillus subtilis:
RT characterization of the arae gene.";
RL J. Bacteriol. 179:1598-1608 (1997).
RN [3]
RP SEQUENCE OF 223-464 FROM N.A., AND FUNCTION.
RC STRAIN=168;
RX MEDLINE=38062200; PubMed=9401028;
RA Sa-Nogueira I., Ramos S.S.;
RT "Cloning, functional analysis, and transcriptional regulation of the
RT Bacillus subtilis arae gene involved in L-arabinose utilization.";
RL J. Bacteriol. 179:7705-7711 (1997).
RN [4]
RP TRANSCRIPTIONAL REGULATION.
RX MEDLINE=99348379; PubMed=10417639;
RA Mota L.J., Favares P., Sa-Nogueira I.M.G.;
RT "Mode of action of AraR, the key regulator of L-arabinose metabolism
RT in Bacillus subtilis.";
RL Mol. Microbiol. 33:476-489 (1999).
CC -1- FUNCTION: Uptake of arabinose across the boundary membrane with
CC the concomitant export of a proton (symport system).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- INDUCTION: Transcription is repressed by the binding of araR to
CC the promoter. L-arabinose acts as an inducer by inhibiting the
CC binding of araR to the DNA, thus allowing expression of the gene.
CC -1- SIMILARITY: Belongs to the sugar transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99121; CAB5401.1; -.
CC EMBL; X98354; CAA66998.1; -.
CC EMBL; Y12105; CAA72812.1; -.
CC PIR; F69587; F69587.
CC Subtilist; BG11907; araE.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1.
```







```
Db 298 QQFTGINAIFVYVPLFSSLSGASGAALN-TVVGVNVGSTMIAVLLSDKFGRRLI 356
Qy 352 ISTGMIIVSLICLGSGLTVAGH-----HPDTKVAWAVALCIASLTSLYIAFFSIGL 401
Db 357 --EGGITCCLAMLAAGITLGVFGQYGTEDLPHPVSAAGVLAV-ICI-----PIAGFAWSW 408
Qy 402 GPITGVYTSEIIPLOVRALGFVAVGNRVTSVISMIFLSLSKAITIGGSPFLYSGIAA 461
Db 409 GPMGNLIPSEITLTETRPAGTAVAVGNFLFSFVIGQAFVSMCLCMKP-GVFLFFAGWLAV 467
Qy 462 VAWVFFTCLETRGRTLEEMGKLFQ-----MPDTGNABEAEADAKEKVE 508
Db 468 INVLCALFLLPETKGVPIERVQALYARHFWKVKVMPAAQEIABDEKRVAAASQAIMK 525
```

Search completed: June 30, 2004, 18:36:01  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:36:04 ; Search time 55 Seconds  
(without alignments)

2942.924 Million cell updates/sec

Title: US-10-051-902A-20

Perfect score: 2559

Sequence: 1 MASDELAKAVEPRKGNVY.....AAEAADAAKVKVLPSSX 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP TREMBL 25:

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mbc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1960.5	76.6	577	Q845E5	Q845E5 oryza sativ
2	1648.5	64.4	511	Q84HC3	Q84HC3 oryza sativ
3	1583	61.9	538	Q84QH3	Q84QH3 prunus cera
4	1515.5	59.2	509	Q84KI7	Q84KI7 prunus cera
5	1509	59.0	523	Q7XA50	Q7XA50 glycine max
6	1493	58.3	481	Q84RI1	Q84RI1 malus domes
7	1492.5	58.3	539	Q9LS92	Q9LS92 arabidopsis
8	1483.5	58.0	545	P93076	P93076 beta vulgar
9	1483.5	58.0	549	P93075	P93075 beta vulgar
10	1482	57.9	519	Q7XB36	Q7XB36 orobanche r
11	1479.5	57.8	511	Q9XIH7	Q9XIH7 arabidopsis
12	1469.5	57.4	491	Q84RI2	Q84RI2 malus domes
13	1463.5	57.2	511	Q9XIH6	Q9XIH6 arabidopsis
14	1420	55.5	513	Q9FOX3	Q9FOX3 apium grave
15	1346	52.6	493	Q23213	Q23213 arabidopsis
16	1316	51.4	508	Q9ZNS0	Q9ZNS0 arabidopsis

17	1274	49.8	506	10	Q8W2W8	Q8W2W8 oryza sativ
18	1274	49.8	506	10	Q7XFD6	Q7XFD6 oryza sativ
19	1274	49.8	574	10	Q9AUM9	Q9AUM9 oryza sativ
20	1272.5	49.7	523	10	Q8RVQ2	Q8RVQ2 apium grave
21	1176	46.0	547	10	Q9SKT9	Q9SKT9 arabidopsis
22	1151	45.0	538	10	Q7XKF1	Q7XKF1 oryza sativ
23	1133.5	44.3	523	10	Q7X6M3	Q7X6M3 oryza sativ
24	1105	43.2	535	10	Q7XKF0	Q7XKF0 oryza sativ
25	668	24.8	457	16	P96742	P96742 bacillus su
26	635.5	24.8	245	10	Q8VZ80	Q8VZ80 arabidopsis
27	629.5	24.6	580	10	Q9C757	Q9C757 arabidopsis
28	624.5	24.4	580	10	Q9C757	Q9C757 rattus norv
29	617	24.1	596	10	Q9XIZ0	Q9XIZ0 oryza sativ
30	615.5	24.1	467	16	Q8CQA7	Q8CQA7 staphylococ
31	613	24.0	409	16	Q83EH4	Q83EH4 coxiella bu
32	610.5	23.9	581	10	Q9LKH1	Q9LKH1 mesembryant
33	610	23.8	544	10	Q93WT7	Q93WT7 olea europa
34	606	23.7	521	10	Q22848	Q22848 arabidopsis
35	602	23.5	546	10	Q93Z41	Q93Z41 arabidopsis
36	600	23.4	509	10	Q8VZR6	Q8VZR6 arabidopsis
37	597	23.3	582	10	Q23492	Q23492 arabidopsis
38	595	23.3	560	10	Q9LFI3	Q9LFI3 arabidopsis
39	594.5	23.2	515	10	Q9LLD8	Q9LLD8 arabidopsis
40	592.5	23.2	534	10	Q9LLE1	Q9LLE1 nicotiana t
41	591.5	23.1	537	3	Q873D5	Q873D5 neurospora
42	589	23.0	555	10	Q9FXV8	Q9FXV8 mesembryant
43	588	23.0	517	10	Q9SX48	Q9SX48 arabidopsis
44	587	22.9	475	10	Q48537	Q48537 prunus arme
45	586	22.9	473	16	Q34718	Q34718 bacillus su

#### ALIGNMENTS

#### RESULT 1

Q945E5 PRELIMINARY; PRT; 577 AA.

AC Q945E5;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Putative sugar transporter.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoidae; Oryzaceae; Oryza.  
CX NCBI\_taxid=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wu P., Chen Q., Huang G., Yi K.;  
RT "Molecular cloning of putative sugar transporter in rice.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
DR EMBL; AF416867; AAL14615.1; -;  
DR Gramene; Q945E5; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005351; F:sugar porter activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0008643; P:carbohydrate transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub. transporter.  
DR InterPro; IPR003663; Sugar transp.  
DR InterPro; IPR005829; Sug transporter.  
DR Pfam; PF00083; sugar tr; 1.  
DR PRINTS; PR00171; SUGTENSPT.  
DR TIGRfams; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR TRANSPORT 1; 1.  
DR PROSITE; PS00217; SUGAR TRANSPORT 2; 1.  
KW Sugar transport; Transmembrane; Transport.  
SQ SEQUENCE 577 AA; 61178 MW; 95651621A5980CD4 CRC64;

Query Match 76.6%; Score 1960.5; DB 10; Length 577;  
 Best Local Similarity 75.8%; Pred. No. 1.9e-125;  
 Matches 368; Conservative 56; Mismatches 65; Indels 3; Gaps 2;

QY 1 MASDELAKEVERKGNVYKASICAILASMASVILGDIYGVNMGAAVYKIDNLTVDQL 60  
 DB 68 MASALAPEAVAPKGNVRFACAILASMTSILGDIYGVNMGASLYKIDFNISDKV 127  
 QY 61 ELILGILSYLFGSPAGARTSDRGRLTVVFAAIVFVGSLLMGFAVNVYVGMAGRFV 120  
 DB 128 EVLMGILNLSYLGSPAGARTSDWGRRTVTVFAAIVFAGFLMGFAVNVYVGMAGRFV 187  
 QY 121 AGVGVYGVNMIAPVYTAETSPASRGFLTTPPEVFINIGILYLSNFAFARPLHLGWR 180  
 DB 188 AGTGVYALMIAPVYTAETSPASRGFLTTPPEVFINIGILYLSNFAFARPLHLGWR 247  
 QY 181 VMLATGAVPSGLLALVFCMPESPRLVILKGLADARAVLEKTSATPEAAERLADIKAA 240  
 DB 248 IMLGGAAPSVLALVLMVGMPSRWLVNMGKGLADAKVLEKTSATPEAAERLADIKAA 307  
 QY 241 AGIPKGLGDVVTVFVKEQGGGELQVWKKLLISPTPAVRRLILSAVGLHFFQOASGDSV 300  
 DB 308 AGIPEELGDVVTVFVKEQGGGELQVWKKLLISPTPAVRRLILSAVGLHFFQOASGDSV 366  
 QY 301 VOYSARLFKSAGITDDNKLGLVTCVAGVTKTFFILVATFLDRAGRRPLLLISGGMIVS 360  
 DB 367 VLYSRVFKSAGITDDNKLGLVTCVAGVTKTFFILVATFLDRAGRRPLLLISGGMIVS 426  
 QY 361 LILCLSGTLVAGHHPDTKVMAVALCIASTLSYIAFFSIGLPGITGVYTSIFFLOVRAL 420  
 DB 427 LILCLAGTLVQGHDPDAKIPWAGISLASTLAYAFFSIGLPGITGVYTSIFFLOVRAL 486  
 QY 421 GPVAVGNRVTVSAVIMTFFLSKAITIGSGFFLYSGIAVAVVFPPTCTLPETRGRTLE 480  
 DB 487 GCSLGVAAVRVTVSAVIMTFFLSKAITIGSGFFLYSGIAVAVVFPPTCTLPETRGRTLE 546  
 QY 481 EMKGLFGPDTGMAREADEAAKKEKVELPSS 512  
 DB 547 ENSKLFQ--DTAAASEDEPAKKEKKEVAAAT 576

RESULT 2  
 Q8LHC3 PRELIMINARY; PRT; 511 AA.  
 AC Q8LHC3  
 DT 01-OCT-2002 (TREMELrel. 22, Created)  
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Similar to myo-inositol transporter 2.  
 GN P0458E05.27.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC clone: P0458E05."  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 DR EMBL: AP004365; BAC05627.1; -.  
 DR Gramene: O8LHC3;  
 DR GO: GO:0036021; C: integral to membrane; IEA.  
 DR GO: GO:0005351; F: sugar porter activity; IEA.  
 DR GO: GO:0005215; F: transporter activity; IEA.  
 DR GO: GO:0008643; P: carbohydrate transport; IEA.  
 DR InterPro: IPR007114; MFS.  
 DR InterPro: IPR005828; Sub transporter.  
 DR InterPro: IPR003663; Sugar\_transpr.

DR InterPro: IPR005829; Sug transporter.  
 DR Pfam: PF00083; sugar tr; 1.  
 DR PRINTS: PR00171; SUGTRNSPORT.  
 DR TIGRFAMs: TIGR00879; SP; 1.  
 DR PROSITE: PS00850; MFS; 1.  
 DR PROSITE: PS00216; SUGAR\_TRANSPORT\_1; 1.  
 DR PROSITE: PS00217; SUGAR\_TRANSPORT\_2; 1.  
 KW Sugar transport; Transmembrane; Transport.  
 SQ SEQUENCE 511 AA; 54606 MW; ECF9573E23E3C625 CRC64;

Query Match 84.4%; Score 1648.5; DB 10; Length 511;  
 Best Local Similarity 88.7%; Pred. No. 3e-104;  
 Matches 334; Conservative 65; Mismatches 70; Indels 17; Gaps 6;

QY 39 IGWNSGAAMVYKIDNLTVDQLEILIGILSYLFGSPAGARTSDRGRLTVVFAAIVF 98  
 DB 19 INVICASLFIKEDMKITDVEVEILLGILNLSLVGSPAGRTSDWIGRRLIILAAVIF 78  
 QY 99 FVGSLLMGFAVNVYVGMAGRFVAGVGVYGVGMIAPIVYTAETSPASRGFLTTPPEVFINI 158  
 DB 79 FVGAIMMGLSVNVYVGMAGRFVAGVGVYGVGMIAPIVYTAETSPASRGFLTTPPEVFINI 138  
 QY 159 GILLGYLSNFAFARPLHLGWRVLMAGVPSGLLALLVFCMPESPRLVILKGLADARA 218  
 DB 139 GILLGVSNYAFSRLEQLQWRMLGVGAAPSVALLALVLMPESPRWLVNMGKGLADAKV 198  
 QY 219 VLEKTSATPEAAERLADIKAAAGIPKGLDGDVVTVFVKEQGGGELQVWKKLLISPTPAV 278  
 DB 199 VLGETSDTAEEAATRLAEIKEAIPAIDLDGDVAVP--KRAGGERRVWKKELISPTPAV 256  
 QY 279 RTLLSAGVLGHFFQOASGDSVVOYSARLFKSAGITDDNKLGLVTCVAGVTKTFFILVAT 338  
 DB 257 RVLLSAGLGHFFQOASGDSVVOYSARLFKSAGITDDNKLGLVTCVAGVTKTFFILVAT 316  
 QY 339 FLDRAGRRPLLLISGGMIVSLICLGSGLTVAGHHPDTKVMAVALCIASTLSYIAFFS 398  
 DB 317 FTLDREGRPLLLASAGGMIAITVLTLGLTLVIGEDA-TGGGWAIASVSIASLAFVAFPS 375  
 QY 399 IGLGPITGVYTSIFFLOVRALGFVAVGNRVTVSAVIMTFFLSKAITIGSGFFLYSG 458  
 DB 376 IGLGPITGVYTSIFFLOVRALGFVAVGNRVTVSAVIMTFFLSKAITIGSGFFLYSG 435  
 QY 459 IAAVAVVFFFTCLPETRGRTLEEMGKLFQMPDTGMAREADEAAK-----EKV---VE 508  
 DB 436 VASLAWLFFFTYLPETRGRTLEQMGELFRIHN--VAGDDSDSATRPPSPPEBEKITYVE 493  
 QY 509 L--PSS 512  
 DB 494 MAAPSS 499

RESULT 3  
 Q84QH3 PRELIMINARY; PRT; 538 AA.  
 AC Q84QH3  
 DT 01-JUN-2003 (TREMELrel. 24, Created)  
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Putative sorbitol transporter.  
 GN SORT2.  
 OS Prunus cerasus.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.  
 OX NCBI\_TaxID=140311;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Montmorency; TISSUE=Fruit;  
 RX MEDLINE=22578918; PubMed=12692316;  
 RA Gao Z., Mauroussat L., Lemoine R., Yoo S.D., Van Nocker S.,  
 RA Loeschner W.;  
 RT "Cloning, Expression, and Characterization of Sorbitol Transporters from Developing Sour Cherry Fruit and Leaf Sink Tissues.";

RL Plant Physiol. 131:1566-1575 (2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Montmorency; TISSUE= Fruit;  
RA Zhifang G., Loescher W.H.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY100638; AAM44082.1; -;  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0005351; F:sugar porter activity; IEA.  
DR GO: GO:0005215; F:transporter activity; IEA.  
DR GO: GO:0008643; P:carbohydrate transport; IEA.  
DR InterPro: IPR007114; MFS.  
DR InterPro: IPR005828; Sub transporter.  
DR InterPro: IPR003663; Sugar transp.  
DR InterPro: IPR005829; Sug transporter.  
DR Pfam: PF00083; sugar tr-1.  
DR PRINTS; PR00171; SUGTRNSPORT.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR TRANSPORT 1; 1.  
DR PROSITE; PS00217; SUGAR TRANSPORT 2; 1.  
SQ SEQUENCE 538 AA; 58569 MW; ADD03B983C6D358 CRC64;  
  
Query Match 61.9%; Score 1583; DB 10; Length 538;  
Best Local Similarity 65.1%; Pred. No. 9.3e-100;  
Matches 310; Conservative 66; Mismatches 96; Indels 4; Gaps 2;  
  
QY 12 PRKGNVYASICAILASMASVILGYDGVMSGAAMYIKKDLNITDQLEILIGILSLY 71  
DB 26 PKPKRNKYAFACALLASMTSILGYDGVMSGAVIYIKKDKVSDVEIVLVGILNLYS 85  
  
QY 72 LFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGYGGMI 131  
DB 86 LIGSAAAGRTSDWIGRRYTIIVLAGAIFPAGALLMGFAPNYAFLMGRFVAGVGYGVALMI 145  
  
QY 132 APVITAEISPAASRGFLTTPEVFNINIGILYLSNFAFARLPHLGWRVMAITGAVPSG 191  
DB 146 APVITAEISPAASRGFLTFPEVFNINIGILYLSNFAFARLPHLGWRVMAITGAVPSI 205  
  
QY 192 LLALVFCWPSRWLVKGLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGV 251  
DB 206 FLATGVLMSPSPRWLVWQGLDARKVLDKTSLSLEESKLRLGEIKAAAGIPKHCNDI 265  
  
QY 252 VVTVPKEGGGELQWKKLILSPTPAVRRILLSAVGLHFFQOAGSDSVVQYSARLFKSA 311  
DB 266 VEVKRSQG--QEVWKQLLRPTPAVRHILMCAVGLHFFQOAGSDAVVLYSPRIEKA 322  
  
QY 312 GITDKNLLGVTCVAGVTKTFPIIVATFLLDRAGRRPRLILSTGMIIVSLICLGSGLTVA 371  
DB 323 GITNPDHVLCTVAGVFKTVFIIVATFLLDRAGRRPRLILSVAGVFTLACLGILITII 382  
  
QY 372 GHPDPTKVAVALCIATSLSYIAFFSIGLGPITGVYTSFIPPLQVRLGFAVGAVNVRV 431  
DB 383 DHSGB-KIMWALSLTWLAYVAFFSIGMGPIITWVYSSEIFFPLQLRAQCGSIGAVNVRV 441  
  
QY 432 TSANVSMTFLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGRITLLEMGKLF 487  
DB 442 VSGVLSMTFISLYKAITIGGAFFLFAA.AAVGWTFFFTMLPETQRTLEDMEVLFG 497

RESULT 4  
Q84KI7 PRELIMINARY; PRT; 509 AA.  
ID Q84KI7  
AC Q84KI7;  
DT 01-JUN-2003 (TREMBLrel. 24, Created)  
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Sorbitol transporter.  
OS Prunus cerasus.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid1; Rosales; Rosaceae; Amygdaloideae; Prunus.  
OX NCBI\_TaxID=140311;

[1]  
SEQUENCE FROM N.A.  
TISSUE= Fruit;  
MEDLINE=22578918; PubMed=12692316;  
Gao Z., Mauroussat L., Lemoine R., Yoo S.-D., van Nocker S.,  
Loescher W.;  
"Cloning, Expression, and Characterization of Sorbitol Transporters  
from Developing Sour Cherry Fruit and Leaf Sink Tissues.";  
Plant Physiol. 131:1566-1575 (2003).  
[2]  
SEQUENCE FROM N.A.  
TISSUE= Fruit;  
Zhifang G., Loescher W.H., Lemoine R.;  
Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
EMBL: AF482011; AAC39267.1; -;  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0005351; F:sugar porter activity; IEA.  
DR GO: GO:0005215; F:transporter activity; IEA.  
DR GO: GO:0008643; P:carbohydrate transport; IEA.  
DR InterPro: IPR007114; MFS.  
DR InterPro: IPR005828; Sub transporter.  
DR InterPro: IPR003663; Sugar transp.  
DR InterPro: IPR005829; Sug transporter.  
DR Pfam: PF00083; sugar tr-1.  
DR PRINTS; PR00171; SUGTRNSPORT.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR TRANSPORT 1; 2.  
DR PROSITE; PS00217; SUGAR TRANSPORT 2; 1.  
SQ SEQUENCE 509 AA; 55407 MW; EA2ED065CE884C8 CRC64;  
  
Query Match 59.2%; Score 1515.5; DB 10; Length 509;  
Best Local Similarity 62.6%; Pred. No. 3.5e-95;  
Matches 299; Conservative 72; Mismatches 102; Indels 5; Gaps 3;  
  
QY 10 VEPKGNVYASICAILASMASVILGYDGVMSGAAMYIKKDLNITDQLEILIGILSL 69  
DB 17 LKPKRNLY- YALGCAILLASMTSILGYDGVMSGASIVIQDKLTSDEVEILIGILNL 75  
  
QY 70 YSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGYGG 129  
DB 76 YSLGSAAGRTSDWIGRRYTIIVAGAIFFTGALLMGLATNYAFLMGRFVAGVGYGVAL 135  
  
QY 130 MIAPVTAETSPASRGFLTTPEVFNINIGILYLSNFAFARLPHLGWRVMAITGAVP 189  
DB 136 MIAPVTAETSPASRGFLTTPEVFNINIGILYLSNFAFARLPHLGWRVMAITGAVP 195  
  
QY 190 SGLLALVFCWPSRWLVKGLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDG 249  
DB 196 SVILAVGLSMSPSPRWLVWQGLGEAKQVLDKTSLSLEEAQLRLADIKAAAGIPHCVE 255  
  
QY 250 DVTVTPKEGGGELQWKKLILSPTPAVRRILLSAVGLHFFQOAGSDSVVQYSARLFK 309  
DB 256 DVTVTPKEGGGELQWKKLILSPTPAVRRILLSAVGLHFFQOAGSDSVVQYSARLFK 312  
  
QY 310 SAGITDKNLLGVTCVAGVTKTFPIIVATFLLDRAGRRPRLILSTGMIIVSLICLGSGLT 369  
DB 313 SAGITDKNLLGVTCVAGVTKTFPIIVATFLLDRAGRRPRLILSTGMIIVSLICLGSGLT 372  
  
QY 370 VAGHPDPTKVAVALCIATSLSYIAFFSIGLGPITGVYTSFIPPLQVRLGFAVGAVN 429  
DB 373 IVDHETE-KMWASVLCITWVLAAYVGFSTGMPFIAMVYSSEIFFPLKLRQAQCGSMGTAVN 431  
  
QY 430 RVTSAVSMTFLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGRITLLEMGKLF 487  
DB 432 RIMSGVLSMTFISLYKAITMGFFLYAGIATVGVVFFFTMLPETQRTLEDMEVLFG 489  
  
RESULT 5  
Q7XA50 PRELIMINARY; PRT; 523 AA.  
ID Q7XA50  
AC Q7XA50;  
DT 01-OCT-2003 (TREMBLrel. 25, Created)



OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX	NCBI_TaxID=3702;
RN	(1)
RP	SEQUENCE FROM N.A.
RC	STRAIN=Columbia;
RA	Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL	Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RT	[2]
OS	Beta vulgaris (Sugar beet).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC	- - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
EMBL	AB026654; BABO1812.1; -
GO	GO:0018021; C:integral to membrane; IEA.
DR	GO:0005351; F:sugar porter activity; IEA.
DR	GO:0005215; F:transporter activity; IEA.
DR	GO:0008643; P:carbohydrate transport; IEA.
DR	InterPro; IPR007114; MFS.
DR	InterPro; IPR005828; Sub transporter.
DR	InterPro; IPR003663; Sugar transp.
DR	InterPro; IPR005829; Sug_transporter.
DR	Pfam; PF00083; sugar_cr; 1.
DR	PRINTS; PR00171; SUGTRANSPORT.
DR	TIGRFAMs; TIGR00879; SP; 1.
DR	PROSITE; PS00850; MFS; 1.
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW	Sugar transport; Transmembrane; Transport.
SQ	SEQUENCE 539 AA; 58102 MW; 36389284E5563A0 CRC64;
	Query Match 58.3%; Score 1492.5; DB 10; Length 539;
	Best Local Similarity 62.0%; Pred. No. 1.4e-93;
	Matches 295; Conservative 76; Mismatches 100; Indels 5; Gaps 4;
OY	11 EPKKNVYKASICAILLASMASVILGYDGVMSGAAMYIKDLNTDVQLLEILGILSLY 70
Db	28 KPFRNN--YAFACALIASMTSILLGYDGVMSGAMVIKDKLDNDUQIGLAGSLNIY 85
OY	71 SLFGSPAGARTSDRIGRRLLTVFAAVFFVGSLMGFAVNVMGLMAGRFVAGVGGYM 130
Db	86 SLTGSCAGRTSDWIGRRITVLGAIIPFAGAILMGLSNPAFLMFGRIAGIVGYALM 145
OY	131 IAPVYTAEISPAASRGFLTTPPEVFINTIGILGYLSNFAPARLPHLGRWRMLAIGVPS 190
Db	146 IAPVYTAEISPASSRGFLNSFPVEVFINAGIMLGYSNLAFSNLPKLVGWRLMGLIGVPS 205
OY	191 GLIALLVFCMPESPRMLVKGLADARAVLEKTSATPEEAERLADIKAAGI PKGLDGD 250
Db	206 VILIAIGVLAMPESPRMLVMQGRIGDAKRVLDKTSPTSEATLRLEDIKHAGIFADCHDD 265
OY	251 VVTVPKGEGGELQVWKKLILSPTPAVRRLILLSAVGLHFHQAGSGSDSVVQY SARLFKS 310
Db	266 VVQV-SERNHGGE-GVWRELLIRPTFAVRVMIAAIGIHFFQQAGSIDAVVLFSPRIEKT 323
OY	311 AGITDDNKLLGVTCVAGVTKTFFILVATHLLDRGRPLLLISTTGMIVSLICIGSLTV 370
Db	324 AGLKTDHQQLLATVAVGWKTSTFILVATFLDRIGRRPLLTSVGVGMVLSAAUGTSLTI 383
OY	371 AGHPDPKAWAVALCISTASTLYIAFSGILGPITGVTTSETIFFPQVRALGFVAVGNR 430
Db	384 I-DOSEKKVMVVVAITATVMTVAIFSIGAGPITWVSISEIFFPLELRSSQGSMGVVNR 442
OY	431 VTSAVISMTFLSKAITIGGSFFLVSGIAAVAVVFFFTCLPETGRITLBEUKGLF 486
Db	443 VTSGVISISFLPNASKAMTTGAFYLFGGIATVANVFFYTFLETQGRMLEDMDELF 498

RESULT 8	
P33076	
ID	P33076 PRELIMINARY; PRT; 545 AA.
AC	P33076;
DT	01-MAY-1997 (TrEMBLrel. 03, Created)
DI	01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	BvCDNA-397 protein.
GN	BvCDNA-397.
OS	Beta vulgaris (Sugar beet).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC	Caryophyllales; Amaranthaceae; Beta.
OX	NCBI_TaxID=161934;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=97198558; PubMed=9046601;
RA	Chou I.-J., Bush D.R.;
RT	Isolation and molecular characteristics of two putative sugar
RT	transporters from sugar beet (Accession Nos. U64902 and U64903)
RT	(PGR97-017).";
RL	Plant Physiol. 113:663-663(1997).
CC	- - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY);
CC	- - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
EMBL	U64903; AMB68029.1; -
DR	GO:0016021; C:integral to membrane; IEA.
DR	GO:0005351; F:sugar porter activity; IEA.
DR	GO:0005215; F:transporter activity; IEA.
DR	GO:0008643; P:carbohydrate transport; IEA.
DR	InterPro; IPR007114; MFS.
DR	InterPro; IPR005828; Sub transporter.
DR	InterPro; IPR003663; Sugar_transpct.
DR	InterPro; IPR005829; Sug_transporter.
DR	Pfam; PF00083; sugar_tr; 1.
DR	PRINTS; PR00171; SUGTRANSPORT.
DR	TIGRFAMs; TIGR00879; SP; 1.
DR	PROSITE; PS00850; MFS; 1.
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW	Sugar transport; Transmembrane; Transport.
SQ	SEQUENCE 545 AA; 59321 MW; 9AD0290DB456BDAB CRC64;
	Query Match 58.0%; Score 1483.5; DB 10; Length 545;
	Best Local Similarity 61.6%; Pred. No. 5.7e-93;
	Matches 299; Conservative 63; Mismatches 112; Indels 11; Gaps 5;
OY	4 DELAKAVPRKNGNVKYASICAILLASMASVILGYDGVMSGAAMYIKDLNTDVQLLEIL 63
Db	25 DPLKK--PPKEN--KFAPACATLASMTSVLLGYDIGVMSGAITYLKDWHISDTQIGVL 79
OY	64 IGILSLYLFSGFACARTSDRIGRRLLTVFAAVFFVGSLLMGFANVMGLMAGRFVAGV 123
Db	80 VGILNYICLFGSFAGRTSDWIGRRITVLGAIIPFVGALLMGATVYAFMWGRFTGTI 139
OY	124 GVYGGMIAPVYTABIISPAASRGFLTTPPEVFINTIGILGYLSNFAPARLPHLGRWRML 183
Db	140 GVGYALTAPVYTAEVSPASSRGFLTSPEVFINAGILGYLSNFLTSPLTHLSWRFML 199
OY	184 ATGAVPSGLLALLVFCMPESPRMLVKORLADARAVLEKTSATPEEAERLADIKAAGI 243
Db	200 GICALPSIFLAIGVLAMPESPRMLVMQGRIGDAKVLNRIISDSPEEAQLRUSEIKQTAGI 259
OY	244 PKGLDGDVTVPGKGGSELQVWKKLILSPTPAVRRLILLSAVGLHFHQAGSGSDSVVQY 303
Db	260 PAECDEDIYKVBKTKKSGN-AWKELFNPTPAVRRAVIAGIGHFFQQAGSIDAVVLY 318
OY	304 SARLFKSAGITDDNKLLGVTCVAGVTKTFFILVATHLLDRAGRPLLLISTTGMIVSLIC 363
Db	319 SPRIFQSAGITNARKQLLATVAVGWKTSTFILVATFLDRIGRRPLLTSYGGMIILT 378
OY	364 LGSGLTVA--GHPTPKAWAVALCIASLTSLVIAFSGILGPITGVTTSETIFFPQVRALG 421
Db	379 LAWSLTVIDHSRH--KITWAIALCITVCVAVASISLGLGPITWVSSEVFPPIRLRAQG 435

```

QY 422 FAVGASNRVTSAVISMTFLSLKAITIGGSPFLYSGIAAVAVWFFFTCLPSTRGRTLEE 481
DB 436 TSMGVAVNRVSVGISIFPLPLSHKITTGAFFLFGGIIAIAWFFFLPSTRGRTLEN 495

QY 482 MGKLF 486
DB 496 MHELF 500

RESULT 9
P93075
AC 293075 PRELIMINARY; PRT; 549 AA.
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BvCDNA-205 protein.
GN BvCDNA-205
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Beta.
OX NCBI_TaxID=161934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97198558; PubMed=9046601;
RA Chlou T.-J., Bush D.R.;
RT "Isolation and molecular characteristics of two putative sugar
RT transporters from sugar beet (Accession Nos. U64902 and U64903)
RT (PGR97-017).";
RL Plant Physiol. 113:663-663 (1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; U64902; AA68028.1; -.
DR FIR; T14606; T14606.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar transport.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR TRANSPORT 1; 1.
DR PROSITE; PS00217; SUGAR TRANSPORT 2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 549 AA; 59790 MW; 895615B299B43A69 CRC64;

Query Match 58.0%; Score 1483.5; DB 10; Length 549;
Best Local Similarity 61.6%; Pred. No. 5.8e-93;
Matches 299; Conservative 63; Mismatches 112; Indels 11; Gaps 5;

QY 4 DELAKAVEPRKGNVYKASICAIIASMASVILGYDGVMSGAAMYIKKDLNITDVQLEIL 63
DB 25 DPLKK---PPKRN--KFAFACATLASMTSVLLGYDGVMSGAIYLEDWHISDTQIGVL 79

QY 64 IGIISLVSLFSGFAGARTSDRIGRRLTVFAAVIFFVGSLLMGFANVYMLMAGRFRVAG 123
DB 80 VGILNIVCLFSGFAAGRTSDWIGRRTIVLAGAIFFFVGALLMGFATNYAFMLVGRFVTGI 139

QY 124 GVGVGGMIAPYVTAIEIPASRGFLTFPEVFINIGILLGVLNFAFARPLHLGWRVML 193
DB 140 GVGVALMIAPYVTAIEVPASRGFLTFPEVFINAGILLGYINLAFSSLPHTLSWRFML 199

QY 194 AIGAVPSGLALLVFCMPSPRWLVKGLADARAVLEKTSATPEEAERLADIKAAGI 243
DB 200 GIGAIPSIFFLAIGLAMPSPRWLMQGLGDAKKVLRNISDSPEEAQLRLSEIKQTAGI 259

```

```

QY 244 PKGLDGDVTPVPGKEQGGELQVWKLLISPTPAVRRIILSAVGLHFFQASGSDSVQY 303
DB 260 PAECDEDIYKVEKTKISGN-AWKELFFNPTPAVRRAVIAGIGIHFFQASGIDAVLY 318

QY 304 SARLFKAGITDDNKLGVTCAGVTKTFILVATFLLDAGRRPLLLISTGMVSLIC 363
DB 319 SPRIFQSAGITNARKQLLATVAVGVWVKTLFILVATFQDKYGRPLLLTSVSGMIITLT 378

QY 364 LGSGLTVA--GHPDPTKVAVAVLACIASLTSYIAFTSIGLGPITGVYTSIPIQVYALG 421
DB 379 LAMSLTVIDHSHH---KITWIALCITMCAVVASFSIGLGPITWYISSEVFPLRUAQG 435

QY 422 FAVGASNRVTSAVISMTFLSLKAITIGGSPFLYSGIAAVAVWFFFTCLPSTRGRTLEE 481
DB 436 TSMGVAVNRVSVGISIFPLPLSHKITTGAFFLFGGIIAIAWFFFLPSTRGRTLEN 495

QY 482 MGKLF 486
DB 496 MHELF 500

RESULT 10
Q7XB36
ID Q7XB36 PRELIMINARY; PRT; 519 AA.
AC Q7XB36;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative mannitol transporter.
GN MAT1.
OS Orobanchaceae.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Orobanchaceae; Orobanchaceae; Orobanchaceae; Orobanchaceae.
OX NCBI_TaxID=46066;
RN [1]
RP SEQUENCE FROM N.A.
RA Delavault P., Simier P., Le Coguic O.;
RT "Identification of a mannitol transporter, OrMat1, in Orobanchaceae."
RT ramosa.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY136668; AA07021.1; -.
DR NCBI_TaxID=46066;
SQ SEQUENCE 519 AA; 56257 MW; 97A54D8D0F20047C CRC64;

Query Match 57.9%; Score 1482; DB 10; Length 519;
Best Local Similarity 58.8%; Pred. No. 6.8e-93;
Matches 294; Conservative 84; Mismatches 112; Indels 10; Gaps 4;

QY 8 KAVEPRKGNVYKASICAIIASMASVILGYDGVMSGAAMYIKKDLNITDVQLEILIGIL 67
DB 20 KALINKPKQM-YALLISWASMTSILLGYDGVMSGATLVIKKDLKITDVQLEILIGLI 78

QY 68 SIYSILFGSPAGARTSDRIGRRLTVFAAVIFFVGSLLMGFANVYMLMAGRFRVAGVGY 127
DB 79 NIYSIGSAGARTSDYLGRRITIVIASVIFPVGAAMGLANNVAVLMVGRFVAGLVGY 138

QY 128 GMIAPVTVTAIEIPASRGFLTFPEVFINIGILLGVLNFAFARPLHLGWRVMLAIGA 187
DB 139 ALMIAPVTAIEVPASRGFLTFPEVFINIGILLGVLNFAFARPLHLGWRVMLAIGA 198

QY 188 VPSGLALLVFCMPSPRWLVKGLADARAVLEKTSATPEEAERLADIKAAGIPKGL 247
DB 199 LPAIFIGLAVIWPSPRWLVNQGSLGDAKKVLRNISDSPEEAQLRLADIKAAGLPEDC 258

QY 248 DGDVTVTPKGGGELQVWKLLISPTPAVRRIILSAVGLHFFQASGSDSVQYSARL 307
DB 259 HDDVTVPLVKQDGGG--GVKELIVHPTKPVHLITIAAVGQFFQASGIDAVVMSPRI 316

QY 308 FKSAGITDDNKLGVTCAGVTKTFILVATFLLDAGRRPLLLISTGMVSLICGSG 367
DB 317 YEKAGITSEKLLATIAVGLCKTVFIVTTFMVDRIGRRVLLTSCGLVLSMLTATG 376

QY 368 LTVAGHHPDPTKVAVAVLACIASLTSYIAFTSIGLGPITGVYTSIPIQVYALG 427

```





Matches 275; Conservative 80; Mismatches 100; Indels 3; Gaps 1;  
QY 30 MASVILGYDGVMSGAAMVKKDLNTDITDVOLEITLILGSLYFSGFAGARTSDRIGRL 89  
Db 1 MSLMGYDLGVMSGASVIEKDKLTQIEIMIGVIEIYSLIGSNAVAGKTSDDWGRY 60  
QY 90 TVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGYGGMIPVYTAISPAASRGELT 149  
Db 61 TVISGAIFGAILMGFSTNVTFLMCGRFVAGVGYALTAPVSAEVSTSRGFLT 120  
QY 150 TPEVFINIGILLYLSNFAFAPLPHLGRVWMLAIGAVPSGLLALLVFCMPESRWLV 209  
Db 121 SPEVEFNIGILLYLSNFAFPCPLDLGLWRLMGLVGAIPSVGLAVGLAMPESRWLV 180  
QY 210 KGRLDARAVLEKTSATPEAAERLADIKAAAGIPKGLDGVVTVPGKEQGGEQVWVK 269  
Db 181 QGRLGAKVLDRTSDSKESMLRLADIKAAAGIPEECNDIVQVSGSHGEG---VWKE 237  
QY 270 LILSPPAVRILLNVGHFTQQAGSDSVVQYARLFKSGAGITDDNKLLGVTCVAVGT 329  
Db 238 LIVHPTTVRHILIAAIGHFQQAGSDIDALVLYSPRVFAKAGITSTNQLLCTVGVGLS 297  
QY 330 KTFILVATFLDRAGRRPLLISGTGMIVSLICLGSGLTVAGHHPDTKVAVAVALCIAS 389  
Db 298 KTVFLVATFLDRVGRRLPLITSWAGWGLVCLGTSITVDQHEGVRMTWAVILCUC 357  
QY 390 TLSYIAFFSIGIPITGVYTSBIFLQVRAALGFAVGVASNRVTSAVISMTFLSKAITI 449  
Db 358 VLAVYGFSSGIPIAVWYSSBIFLRLRAQCGGVAVNRLMSGILSMTISLYKAITM 417  
QY 450 GGSFELYSGIAVAVVFFTCIPETRGRTLEBMKGLFG 487  
Db 418 GGTFFLYAAGTVGMIFFFTNLPTQGRRTLEDMEVLF 455

## RESULT 13

Q9XIH6 PRELIMINARY; PRT; 511 AA.  
ID Q9XIH6  
AC Q9XIH6  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative sugar transporter.  
GN ATG16130.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Sree T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
RA Buell C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
thaliana.";  
RL Nature 402:761-768(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
DR EMBL; AC007134; AAD26955.1; -.  
DR PIR; A84537; A84537.  
DR GO; GO:0016021; C:integral to membrane; IEA.

GO; GO:0005351; F:sugar porter activity; IEA.  
GO; GO:0005215; F:transporter activity; IEA.  
GO; GO:0008643; P:carbohydrate transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub.transporter.  
DR InterPro; IPR003663; Sugar transport.  
DR InterPro; IPR005829; Sug transporter.  
DR Pfam; PF00083; sugar tr; 1.  
DR PRINTS; PR00171; SUGTRANSORT.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Sugar transport; Transmembrane; Transport.  
SQ SEQUENCE 511 AA; 54932 MW; 60CB43A9DB5C8396 CRC64;  
Query Match 57.2%; Score 1463.5; DB 10; Length 511;  
Best Local Similarity 58.4%; Pred. NO. 1.2e-91;  
Matches 290; Conservative 73; Mismatches 129; Indels 5; Gaps 2;  
QY 11 EPRKKGKVKVASICAILASMASVILGYDGVMSGAAMVKKDLNITDVOLEILIGLSLY 70  
Db 16 EPRGNRSRPAFACAILASMTSILGYDGVMSGAALFIKDDLKSLDVOLEILMGLNIY 75  
QY 71 SLFGSPAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGYGGM 130  
Db 76 SLIGSAGARTSDWIGRRYTVILAGFFFCGALLMGFATNYPFIMVGRFVAGVGYAMM 135  
QY 131 IAPVYTAISPAASRGFLTTEPEVFINIGILLYLSNFAFAPLPHLGRVWMLAIGAVPS 190  
Db 136 IAPVYTEVAPASRGFLSSPEIFINIGILLYVSNYFFAKLPEHIGRFLMGLGAVPS 195  
QY 191 GLLALLVFCMPESRWLVKGRLDARAVLEKTSATPEAAERLADIKAAAGIPKGLDGD 250  
Db 196 VFLAIGVLA MPESRWLVNQGRIGDAFKVLDKTSNTKEAISRLNDIKRAVGIPDDMTDD 255  
QY 251 VTVPGKEQGGEQVWVKLLISLTPPAVRILLNVGHFTQQAGSDSVVQYARLFK 310  
Db 256 VIVPNKKSAG--KGVKDLLVRPTPSVRHILIAAGLGHFSQAGSIDAVVYSPTFSR 313  
QY 311 AGITDDNKLLGVTCVAVGTCTFFILVATELDRAGRRPLLISGTGMIVSLICLGSGLTV 370  
Db 314 AGLSKNDQLLATVAVGVTCLFVIVGTCLVDRFGRALLTSMGMFFSLTALGTSLTV 373  
QY 371 AGHPDTKVAVAVLALCIASLTSLYIAPFSICLGPITGVYTSBIFLQVRAALGFAVGNR 430  
Db 374 IDNPGQTLKWAIGLAVTTMTFVATFSLGAGPVTWVASEIFPVRLRAQASLGVMNLN 433  
QY 431 VTSAVISMTFLSKAITTIGSGFFLYSGIAVAVVFFTCIPETRGRTLEEMKGLFGMPD 490  
Db 434 LMSGIIGMTFLSKGLTIGGAFLLPAGVAVAAWVFFFTLPETRGVPLEBIESLFG--- 490  
QY 491 TGMABEAEADAAAEKVV 507  
Db 491 SYSANKKNVMSGKQV 507  
RESULT 14  
Q9FOX3 PRELIMINARY; PRT; 513 AA.  
ID Q9FOX3  
AC Q9FOX3;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Mannitol transporter.  
GN Manl.  
OS Apium graveolens var. dulce.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Campanulids; Apiales; Apiaceae; Apioideae; apioid superclade;  
OC Apium clade; Apium.  
OX NCBI\_TaxID=117781;  
RN [1]

```
RP SEQUENCE FROM N.A.
RC TISSUE=Phloem;
RX MEDLINE=21149881; PubMed=11251106;
RA Noiraud N., Maurousset L., Lemoine R.;
RT "Identification of a mannitol transporter, agmat1, in celery phloem.";
RL Plant Cell 13:695-705(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF215837; AAC33998.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005351; F: sugar porter activity; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0008643; P: carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub. transporter.
DR InterPro; IPR003663; Sugar transp.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00083; sugar tr_1.
DR PRINTS; PR00171; SUGTRNSPRT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 513 AA; 56013 MW; E08D6555B9B08781 CRC64;

Query Match 55.5%; Score 1420; DB 10; Length 513;
Best Local Similarity 55.6%; Pred. No. 1.1e-88;
Matches 282; Conservative 91; Mismatches 116; Indels 18; Gaps 6;

QY 11 EPRKGNVYKASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQLEILIGILSLY 70
DB 15 KPKPKRN-KYAFACALLASMSILGYDTGVLGSASIVYKEDLHFSVDQVEIIGIINIY 73

QY 71 SLFGSPAGATSDRIGRLTVVFAAVIFFVGLSLLMGFAVNYGMLMAGRFVAGVGVGGM 130
DB 74 SLIGSAIAGRTSDMIGRRYTWLAGIIFPLGAIFMGLATNFAFLMFGRFVAGIGVGYAMM 133

QY 131 IAPVYTAISPAASRGFLTTPEVFINIGILLVLSNFAFARLPLHLGWRVMLAIGAVPS 190
DB 134 IAPVYTAEVAPSSRGFLTSPPEVFINSVGLLVYVSNFAFAKCFPLWGLWRMLGIGAPPS 193

QY 191 GILLALLVFCMPSPRWLVKGRADARAVLEKTSATPEEAERLADIKAAAGIPKGLGD 250
DB 194 VALAIIVLYMPSPRWLVNQGLGEARVLEKTSKEEAHQRLSDIKERAGIDKDCND 253

QY 251 VVTPGKEGGGELQVWKKLILSPPAVRRILLGAVLHFTQQASGSDSVVQYSARLPKS 310
DB 254 VVQVPEKTK--DEAVWKEILIHPTKPVVRAAITGIGHFTQQACGIDAVVLYSPRIFEK 310

QY 311 AGITDDNKLGVTCVAGVYTKTFFILVATFLDRAGRRPLLLISTGGMIVSLICLGSLTV 370
DB 311 AGIKNSKKLLATIAVGCKTVLILSTFQDKIGRRPLMLTSMGGVIALFVLAGSLTV 370

QY 371 --AGHPDPTKVAMAVALCIASLTSLVAFPSIGLGPITGVYTSSEIFPLQVRLGFAVGAS 428
DB 371 INKSHHTG---HWAGGLAIFTVYAFVSIFSSGMGFIAMVYSSEVFPPLRLAQGCSIGVAV 427

QY 429 NPVTSVAVTSMFLSLSKAITTIGGSFFLYSGIAAVAWVFFTCLPETRGRTLEEMGKFG- 487
DB 428 NRGMSGIIGMTFISMYKAWTIGGAFLLFAVVASIGWVFMTFPTQGRNLEERIELLFGS 487

QY 488 -----MPDTGMABEADAARKEVW 507
DB 488 YFGWRKTLKDL-KAKEAAEAKSRESEV 513

RESULT 15
O23213
ID O23213 PRELIMINARY; PRT; 493 AA.
AC O23213;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter like protein.
GN At4G36670.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RA Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Schueller C.,
RA Chawatzi N.;
RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; Z99708; CAB16808.1; -.
DR EMBL; AL161589; CAB80333.1; -.
DR PIR; A85433; A85433.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005351; F: sugar porter activity; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0008643; P: carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub. transporter.
DR InterPro; IPR003663; Sugar transp.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPRT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR Sugar transport; Transmembrane; Transport.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 493 AA; 52922 MW; E49D9E7298E9982A CRC64;

Query Match 52.6%; Score 1346; DB 10; Length 493;
Best Local Similarity 55.5%; Pred. No. 1.2e-83;
Matches 269; Conservative 94; Mismatches 115; Indels 6; Gaps 5;

QY 6 LAKAVEPRKGNV-KYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQLEILI 64
DB 1 MADQIGKEPAGVNRREALCAIVASIVSIIFGYDTGVMSCAMVFIEDLKTNDVQLEVL 60

QY 65 GILSLYSLFGSPAGATSDRIGRLTVVFAAVIFFVGLSLLMGFAVNYGMLMAGRFVAGV 124
DB 61 GILNLCALVGLSLAGRTSDIIGRYTIVLASILFMLGSLMGMPNYPVLLSGRCTAGLG 120

QY 125 VGYGGMIAPIVYTAISPAASRGFLTTPEVFINIGILLVLSNFAFARLPLHLGWRVMLA 184
DB 121 VGFALVAVPVYSAEIATASHRGLASLPHLCISIGILLGVYVNYFSSKLPMLHIGWRLMG 180

QY 185 IGAVPSGLIALLVFCMPSPRWLVKGRADARAVLEKTSATPEEAERLADIKAAAGI- 243
DB 181 IAAVPSLVLAFLGFLKMPSPRWLMQGRLEKEILELVSNSPEEAELRFQD;KAAAGID 240

QY 244 PKGLGDGVTVTPGKEGGELQVWKKLILSPTPAVRRILLGAVLHFTQQASGSDSVVQY 303
DB 241 PKVD-DVVMEKKTHTGE--GWKELILRPTPAVRVLLTALGIHFQHASIEAVLLY 297

QY 304 SARLPKSAGITDDNKLGVTCVAGVYTKTFFILVATFLDRAGRRPLLLISTGGMIVSLIC 363
DB 298 GPRIFKAGITTKDKLFLVTIGVIMKTTFIPTATLLDKVGRKKLLTSVGGMVIALTM 357

QY 364 LGSGLTVAGHPDPTKVAMAVALCIASLTSLVAFPSIGLGPITGVYTSSEIFPLQVRLGFA 423
DB 358 LGFGLTWA-QNAGKLAVALVLSIVAYSFVAFPSIGLGPITVWYSEVFPPLKRAQGAS 416

QY 424 VGVASNRVTSVAVTSMFLSLSKAITTIGGSFFLYSGIAAVAWVFFTCLPETRGRTLEEMG 483
DB 417 LGVAVNRVNTVSMFLSLTSITTTGGAFFMFAGVAAVAVNFFFLPETKGSLEEIE 476
```

QY 484 KLF 486  
Db 477 ALF 479

Search completed: June 30, 2004, 18:42:52  
Job time : 69 secs